Stuttering:
Studies of therapy outcome and speech motor control

Print: Printpartners Ipskamp, Enschede
Cover design: W.J. Huinck
Lay-out: D. Helsper & W.J. Huinck

ISBN-10: 90-9021012-1
ISBN-13: 978-90-9021012-4

Stuttering: Studies of therapy outcome and speech motor control.
Thesis Radboud University Nijmegen Medical Centre, Nijmegen.

Stuttering:
Studies of therapy outcome and speech motor control

Een wetenschappelijke proeve op het gebied van de Medische Wetenschappen

PROEFSCHRIFT

ter verkrijging van de graad van doctor
aan de Radboud Universiteit Nijmegen
op gezag van de Rector Magnificus, prof. dr. C.W.P.M. Blom,
volgens besluit van het College van Decanen
in het openbaar te verdedigen op
maandag 13 november 2006
om 13.30 uur precies

door

Wendy Josine Huinck

geboren op 27 april 1970
te Maastricht

Promotores:	Prof. dr. K. Graamans
	Prof. dr. W. Hulstijn
Copromotores:	Dr. H.F.M. Peters
	Dr. B.A.M. Maassen
Manuscriptcommissie:	Prof. dr. F.W. Kraaimaat (voorzitter)
	Prof. dr. A.C.M. Rietveld
	Prof. dr. J. Van Borsel (Universiteit Gent)

Aan mijn ouders

Voor Rogier,
Sebastiaan & Semuel

CONTENTS

Chapter One

Introduction and thesis outline

General Introduction

Stuttering is a communication disorder in which the flow of speech is interrupted by repetitions (stu-stu-stu-stuttering), prolongations (ssssssttuttering), or abnormal stoppages or blockades (s…….tuttering) of sounds and syllables. In addition to these dysfluencies, persons who stutter often experience physical tension and struggle in their speech muscles (like unusual facial and body movements associated with the effort to speak), as well as embarrassment, anxiety, and fear about speaking (*e.g.* Wingate, 1964). Although there are many different patterns of stuttering and many degrees, ranging from mild to severe stuttering, all persons who stutter have in common that the symptoms of stuttering impair speaking and effective communication.

The exact cause of stuttering is unknown. Researchers around the world are actively seeking new information about this complex communication disorder. Factors that seem to contribute to the development of stuttering are: (1) genetics (about 50% of those who stutter have a family member who does also, *e.g.* Yairi, Ambrose, & Cox, 1996); (2) development during childhood (children with other speech and language problems or developmental delays are more likely to stutter, see Nippold, 1990 for a review); (3) neurophysiology (*e.g.* De Nil, Kroll, Kapur, & Houle, 2000 provided qualified support for the hypothesis that stuttering adults show atypical lateralization of language processes); (4) and family dynamics (high expectations and fast-paced lifestyles can contribute to stuttering). Approximately 1% of the population stutters, with four times as many males as females (*e.g.* van Riper, 1982).

There is no cure for stuttering, but therapy programmes have been developed in which persons who stutter learn to speak more easily and more fluently, to feel better about themselves and their speaking ability, or to communicate more effectively. The efficacy of these therapy programmes is widely discussed. Specifically with the launch of 'evidence-based medicine' (EBM) and 'evidence-based practice' (EBP) (Sacket, Rosenberg, Gray, Haynes, Richardson, 1996), this topic gained much interest (*e.g.* Bernstein Ratner, 2005; Bernstein Ratner & Healey, 1999; Bothe, 2003; Conture, 1996, 2001; Craig, 2002; Eichstaedt, Watt & Girson, 1998; Finn, 2003ab; Huinck & Donders 2005; Ingham, 2003; Langevin & Kully 2003; Onslow, 2003; Yaruss, 1998ab; Yaruss, 2001; Yaruss & Quesal, 2004). Researchers and clinicians agree on the need of treatment outcome studies, but, at the same time, they disagree about how recovery should be defined (Finn, Howard & Kubala, 2005). Some of them are of the opinion that the reduction and elimination of stuttering and its replacement with natural sounding, stutter-free speech in everyday settings is the minimal standard for defining recovery (Ingham & Cordes, 1997), while others assert that the elimination of stuttering is not possible because reactions to stuttering, such as fear and avoidance, are the most fatiguing features of the disorder. Only if these problems are decreased, a full recovery might be possible (Manning, 2001). So, standardization with respect to therapy goals -although probably a challenging task- is needed. Especially in adults there are some complicating factors in efficacy research, for example: (1) The heterogeneity in the group of stuttering adults. Adults may have a history of negative and/or positive experiences with their dysfluency, which directly influences the type, place and severity of the dysfluency. (2) The cause of stuttering is still unclear (as mentioned above). Some researchers focus on the brain (see Ingham, 2001 and de Nil, 2004 for a review), on speech motor control (see Van Lieshout, Hulstijn & Peters, 2005 for a review), psycholinguistics (see Conture, Courtney, Zackheim, Anderson & Pellowski, 2004 for a review), or genetics (*e.g.* Felsenfeld, 1997; Yairi & Ambrose, 1996; Felsenfeld, Kirk, Zhu, Statham, Neale, & Martin, 2000). Recently, researchers are also open for a multi-factor approach (see Van Lieshout, Hulstijn & Peters, 2004), but the fact that there is no pointed cause of stuttering complicates the development of therapy programmes. (3) Within the same individual, the level of fluency varies, depending on the day, the emotions, the mastery of a given conversation subject and the different communication situations within everyday life (*e.g.* de Andrada, Cervone, Sassi, 2003). (4) The use of fluency enhancing techniques often results in unnatural and tensed speech, so-called pseudofluency (*e.g.* Ingham, Gow & Costello, 1985; Kalinowski, Noble, Armson & Stuart, 1994; Dayalu & Kalinowski, 2001, 2002) and thus, although maybe fluent after treatment, the speech might sometimes still be deviant from the speech of non-stuttering people. (5) Clients often relapse at the long term after therapy (*e.g.* Boberg, 1981; Bray, kehle, Lawless & Theodore 2003; Craig, 1998; Eichstadt, Watt & Girson, 1998; Finn 2003ab, Ryan & van Kirk Ryan, 1995). Therefore, long-term measurements are required when treatment efficacy is investigated.

Thus, whereas researchers seek to improve their knowledge about the key factors that influence therapy success, they still lack information about the exact mechanisms that underlie the complex phenomenon of stuttering. By studying the effectiveness of therapy (studying what levels or aspects of speech production are changed and improved by the different treatment programs) and by investigating how persons who stutter differ from persons who do not stutter, we aim to improve insights into therapy effects. Based on how persons who stutter differ from persons who do not stutter, therapy goals can be formulated. In addition, long-term treatment results reveal information about the specific types of stuttering behaviour that can be changed and about what types appear to be more persistent. This can help us better understand stuttering.

Thesis outline

This thesis is divided in two parts. Part I addresses stuttering therapy efficacy. In the first Chapter of this part (Chapter 2), a longitudinal study on the effect of three different stuttering therapy programmes is described: The Comprehensive stuttering program (CSP), Individualised stuttering therapy according to guidelines formalised by the Association of Stuttering Treatment Centres in the Netherlands (VSN, Vereniging Stottercentra Nederland), and the Doetinchemse Method (DM). The aim was to evaluate each of these therapy programmes separately in order to improve our understanding of changes at different levels of stuttering, e.g. fluency, speech rate, negative or positive cognitions and emotions, attitudes towards stuttering and also judgments of listeners about the speech. Specifically in the light of the ongoing discussion on long term efficacy of stuttering therapies, long-term measures (one and two years post treatment) were taken.

In the following two Chapters of this thesis, treatment effects of the CSP in different groups of stuttering people are described. Chapter 3 concentrates on the relationship between pre-treatment clinical profiles and the actual treatment outcome. A better understanding of how different profiles of stuttering contribute to treatment outcome might help us to improve the selection of treatment strategies or programmes. Chapter 4 addresses cross-cultural differences. The therapy results of the CSP in a Canadian group are compared with those of a Dutch group. The CSP programme is developed in Canada and is introduced in the Netherlands in 2000. Although the CSP has shown to be fairly effective in Canada (Langevin & Boberg, 1993) and in the Netherlands (Huinck & Peters, 2004, see Chapter 2), it was not known if cultural differences caused differences in treatment outcome. This study was set up to empirically investigate whether treatment effectiveness of the CSP can be generalized from one country (or culture) to another. That is, a treatment programme must be sensitive to the values and attitudes of the culture or subculture in which the treatment is being delivered (see Taylor, 1986).

In Chapter 5 we tested the validity of a simple outcome measure to assess the efficacy of stuttering therapies. Reason for this is that in the field of stuttering therapy, a large set of tests (*e.g.* self-evaluation questionnaires, methods for scoring the percentage of dysfluency and speech rate) has been developed not only for diagnostic procedures but also for the evaluation of post or within treatment efficacy. Most of these tests are rather time-consuming. The validity of a simple and not time-consuming self-assessment scale was tested by relating it to objective measures and other (self-) evaluation tests.

The findings of Chapter 3, where it is shown that different stuttering profiles result in different therapy outcomes, raise the question of how to characterise degrees in stuttering severity: What are the underlying mechanisms that are at fault? Or, in other words, what mechanisms cause a staggering speech motor control in stuttering persons but do not in non-stuttering speakers? The second part of this thesis describes studies that focus on differences between PWS and PWNS. Contrary to part one we will not focus on aspects of emotion and cognition related to stuttering, but we will limit ourselves to aspects of speech motor control. Measures that tap into speech production processes like speech reaction times and word durations are applied to indicate differences between the two groups. To this end, subscales of 'The Nijmegen Speech Motor Test' (NMST) are analysed. The NSMT is developed (and is still under construction) to classify participants into specific stuttering profiles, based on aspects of speech motor control. Part II of this thesis describes the first results of several speech tasks of the NSMT. Better understanding of stuttering profiles results in better therapy choices, resulting in improved treatment outcomes. In other words, in the first part of this thesis (Chapter 3), classification of stuttering was based on stuttering symptoms on the surface of stuttering (*e.g.* speech rate, percentage stuttered syllables, struggle behaviour). With the NSMT, *underlying* aspects of stuttering were measured, using a simultaneous registration of phonation, articulation and respiration in a set of speech tasks.

In Chapter 6, a study is described in which it is investigated if persons who stutter differ from persons who do not stutter in the co-production of different types of consonant clusters. The main goal was to find out if articulatory complexity has influence on stuttering, to improve our understanding of stuttering at the level of speech execution. Diadochokinesis, sentence repetition and reaction time tasks and variability in both groups is investigated in Chapter 7. The thesis ends with a general summary and a discussion (Chapter 8).

References

Bernstein Ratner, N., & Healey, E.C. (1999). Bridging the gap between stuttering research and practice: An overview. In N. Bernstein Ratner & E.C. Healey (Eds.), *Stuttering research and practice: Bridging the gap* (pp. 1-12). Mahwah, New Jersey: Lawrence, New Jersey.

Bernstein Ratner, N. (2005). Evidence-based practice in stuttering: Some questions to consider. *Journal of Fluency Disorders, 30,* 163-188.

Boberg, E. (1981). Maintenance of fluency: An experimental program. In E. Boberg (Ed.), New York: Elsevier North-Holland.

Bothe, A.K. (2003). Evidence-based treatment of stuttering: V. The art of clinical practice and the future of clinical research. *Journal of Fluency Disorders, 28,* 247-258.

Bray, M.A., Kehle, T.J., Lawless, K.A., & Theodore, L.A. (2003). The relationship of self-efficacy and depression to stuttering. *American Journal of Speech-Language Pathology, 12,* 425-431.

Conture, E.G. (1996). Treatment efficacy: Stuttering. *Journal of Speech and Hearing Research, 39,* 518-526.

Conture, E.G. (2001). *Stuttering: Its nature, diagnosis, and treatment.* Boston: Allyn & Bacon.

Conture, E.G., Zackheim, C.T., Anderson, J.D., & Pellowski, M.W. (2004). Linguistic processes and childhood stuttering: many's a slip between intention and lip. In B. Maassen, R.D. Kent, H.F.M. Peters, P.H.H.M. Van Lieshout, & W. Hulstijn (Eds.), *Speech motor control in normal and disordered speech* (pp. 253-281). Oxford: University Press.

Craig, A. (1998). Relapse following treatment for stuttering: a critical review and correlative data. *Journal of Fluency Disorders, 23,* 1-30.

Craig, A. (2002). Fluency outcomes following treatment for those who stutter. *Perceptual and Motor Skills, 94,* 772-774.

Dayalu, V.N., & Kalinowski, J. (2001). Re: Stuttering therapy results in pseudofluency. *International Journal of Language & Communication Disorders, 36,* 405-408.

Dayalu, V.N., & Kalinowski, J. (2002). Pseudofluency in adults who stutter: the illusory outcome of therapy. *Perceptual and Motor Skills, 94,* 96.

De Andrade, C.R.F., Cervone, L. M., & Sassi, F.C. (2003). Relationship between the stuttering severity index and speech rate. *Sao Paulo Medical Journal, 121,* 81-84.

De Nil, L.F., Kroll, R.M., Kapur, S., & Houle, S. (2000). A positron emission tomography study of silent and oral single word reading in stuttering and nonstuttering adults. *Journal of Speech, Language, and Hearing Research, 43*, 1038-1053.

Eichstaedt, A., Watt, N., & Girson, J. (1998). Evaluation of the efficacy of a stutter modification program with particular reference to two new measures of secondary behaviors and control of stuttering. *Journal of Fluency Disorders, 23,* 231-246.

Felsenfeld, S., Kirk, K.M., Zhu, G., Statham, D.J., Neale, M.C., & Martin, N.G. (2000). A study of the genetic and environmental etiology of stuttering in a selected twin sample. *Behavior Genetics, 30,* 359-366.

Finn, P. (2003). Evidence-based treatment of stuttering: II. Clinical significance of behavioral stuttering treatments. *Journal of Fluency Disorders, 28,* 209-218.

Finn, P. (2003). Addressing generalization and maintenance of stuttering treatment in the schools: a critical look. *Journal of Communication Disorders, 36,* 153-164.

Huinck, W.J., & Donders, A.R.T. (2005). Meten van effectiviteit van stottertherapie. In H.F.M. Peters, R. Bastiaanse, J. Van Borsel, Ph. Dejonckere, K. Jansonius-Schultheiss, S. van der Meulen, and B.J.E. Mondelaers (Eds.), Handboek Stem-, Spraak-, en Taalpathologie [B7.3.3], (pp 1-18). Houten: Bohn Stafleu Van Loghum.

Ingham, J.C. (2003). Evidence-based treatment of stuttering: I. Definition and application. *Journal of Fluency Disorders, 28,* 197-207.

Ingham, R.J., Gow, M., & Costello, J.M. (1985). Stuttering and speech naturalness: Some additional data. *Journal of Speech and Hearing Disorders, 50,* 217-219.

Kalinowski, J., Noble N., Armson, J., & Stuart, A. (1994). Naturalness ratings of the pretreatment and post-treatment speech of adults with mild and severe stuttering. *American Journal of Speech-Language Pathology, 3,* 61-66.

Langevin, M., & Kully, D. (2003). Evidence-based treatment of stuttering: III. Evidence based practice in a clinical setting. *Journal of Fluency Disorders, 28,* 219-236.

Nippold, M.A. (1990). Concomitant speech and language disorders in stuttering children. A critique of the literature. *Journal of Speech and Hearing Disorders, 55*, 51-60.

Onslow, M. (2003). Evidence-based treatment of stuttering: IV. Empowerment through evidence-based treatment practices. *Journal of Fluency Disorders, 28,* 237-245.

Ryan, B.P., & Van Kirk Ryan, B. (1995). Programmed stuttering treatment for children: comparison of two establishment programs through transfer, maintenance and follow up. *Journal of Speech and Hearing Research, 38,* 75.

Sacket, D.L., Rosenberg, W.M.C., Gray, J.A.M., Haynes, R.B., & Richardson, W.S. (1996). Evidence based medicine: what it is and what it isn't. British medical journal [On-line]. Available: http://cebm.jr2.ox.ac.uk/ebmisisnt.htm.

Van Riper, C. (1982). The Nature of stuttering. Englewood Cliffs: Prentice-Hall,inc.

Wingate, M.E. (1964). A standard definition of stuttering. *Journal of speech and hearing disorders, 11,* 484-489.

Yairi, E., Ambrose, N.G., & Cox, N.J. (1996). Genetics of stuttering: a critical review. *Journal of Speech, Language, and Hearing Research, 39,* 771-784.

Yaruss, J.S. (1998). Treatment Outcomes in Stuttering: Finding Value in Clinical Data. In A.K. Cordes & R.J. Ingham (Eds.), *Treatment Efficacy for Stuttering. A Search for Empirical Bases* (pp. 213-243). San Diego: Singular Publishing Group, Inc.

Yaruss, J.S. (1998). Describing the consequences of Disorders: Stuttering and the International Classification of Impairments, Disabilities, and Handicaps. *Journal of Speech, Language, and Hearing Research, 41,* 249-257.

Yaruss, J.S. (2001). Evaluating treatment outcomes for adults who stutter. *Journal of Communication Disorders, 34,* 163-182.

Yaruss, J.S., & Quesal, R.W. (2004). Stuttering and the international classification of functioning, disability and health (ICF): an update. *Journal of Communication Disorders, 37,* 35-52.

Part I

Stuttering therapy efficacy

Chapter Two

Efficacy Research in Stuttering Therapy
A longitudinal observation of the effects of three treatment programmes

A slightly adapted version of:
Wendy J. Huinck & Herman F.M. Peters (2004). Nijmegen University Press.
Funded and approved by the Dutch Health Insurance Board (CvZ, College van Zorgverzekeringen).

Parts of this chapter have been published as research report nr 25003735 by the Dutch Health Insurance Board, and is april 28th 2005 presented to the ministry of Health, Welfare and Sport.

1 Introduction

1.1 Motivation of the study

Stuttering is a complex speech disorder that is characterised by an abnormally high frequency or duration of dysfluencies in the forward flow of speech. These dysfluencies usually manifest themselves in (a) repetitions of sounds, syllables, or one-syllable words, (b) prolongations of sounds, or (c) 'blocks'[1] of airflow or voicing in speech. Individuals who stutter are usually aware of their speech impediment and are often embarrassed by it. Moreover, they tend to use excessive physical and mental effort to speak (Guitar, 1998).

Three to five percent of all children experience brief periods of stuttering during their childhood. Of these children, 50 to 80% recover spontaneously (Yairi & Ambrose, 1999) and, although the majority of the remaining children recover after therapy, some become persistent stutterers. With a prevalence of approximately 1%, stuttering can be said to be a relatively common speech problem. Besides its high prevalence, stuttering is also considered one of the most perplexing problems within the field of communication disorders (Ansel, 1993) and in many cases it has a considerable impact on the stutterer's quality of life.

Many theories about the aetiology of stuttering have been formulated and its origin has been attributed to areas as far apart as genetics and cognitive dysfunction. Some researchers have found relative slowness of laryngeal, respiratory and articulatory reactions times in stutterers (*e.g.* Watson & Alfonso 1982, 1983, 1987; Van Lieshout, Hulstijn, & Peters 1996a; Huinck, Wouters, Hulstijn & Peters, 2001) and effects of auditory delayed feedback on the fluency of stuttering and nonstuttering speakers (*e.g.* Kalinowski, Stuart, Sark, & Armson, 1996). Others have reported on the involvement of neurological and cognitive processes (*e.g.* De Nil, 1999; De Nil & Bosshardt 2001) and genetic factors to explain stuttering (*e.g.* Yairi, Ambrose, & Cox, 1996). The various theories all exert their influence on the approach of and methods applied in clinical interventions.

In recent years, a number of different stuttering therapies have been developed and, although they have a certain overlap, they clearly differ in their approach, treatment goals and speech remediation techniques. To some extent, this diversity is related to the fact that there is still some controversy about the appropriate goals for stuttering therapy (Starkweather, 1993). The clinician's beliefs about the cause and development of stuttering need to be compatible with the assessment strategies, therapy goals and therapy procedures (Guitar, 1998). Two current approaches can be distinguished: (1) 'stuttering modification therapy' that teaches the stutterer to modify his moments of stuttering, among other techniques by reducing the tension and (2) 'fluency shaping therapy' that helps the stutterer to speak more fluently, for example by altering his speech pattern (Guitar, 1998). Furthermore, therapy can be based on an individual or group approach.

In addition to the issues concerning treatment goals, there is an ongoing debate about treatment efficacy. It is suggested in the literature that current stuttering therapies are indeed effective: "…that substantial improvement, as defined in these studies, typically occurs as a result of almost any kind of therapy in about 60 to 80

[1] *The term 'block' is currently exclusively used to denote stuttering behavior in which the speaker stops the flow of air or voice. This differs from its historical usage when it referred to any expression of stuttering.*

percent of the cases" (Bloodstein, 1995, p. 438). However, Bloodstein probably overestimates the actual success of stuttering treatment in that the few evaluation studies he refers to were not carried out by independent investigators and they, moreover, were mostly limited to the proportion of dysfluencies and/or post-treatment speech rate (Franken, 1995). It is clear that measurements of proportional dysfluencies and speech rate yield an incomplete picture of the effects of treatment on stuttering.

Despite the numerous therapies that have been developed, there has been little impartial, scientific research that has focused on the efficacy of these treatment programmes. Then again, given the complexity of stuttering and the diversity in treatment approaches, this is not surprising.

There is consensus, however, that stuttering therapies should comply with benchmarks derived from evidence-based medicine and/or evidence-based practice (Sacket, Rosenberg, Gray, Haynes, & Richardson, 1996), which makes treatment-outcome studies in the field of stuttering indispensable. To this end, in January 1999 a longitudinal study[2] was initiated at the University Medical Centre St Radboud, Nijmegen, the Netherlands, which was completed at the end of 2003. The aim of the study was to assess three different stuttering therapy programmes as to their impact on stuttering-related speech fluency, speech characteristics, speech movements, emotions and cognitions. The rationale of the research project will be discussed next.

1.2 Aim of the research project

The present longitudinal study evaluates the efficacy of three stuttering therapy programmes. The programmes all differed in treatment method, approach and goals. To assess their efficacy overt speech behaviours (speech fluency, speech quality, and speech-motor control) and subjective indices were analysed by means of self-evaluations of the cognitive and emotional aspects related to stuttering. The therapies were examined separately on both their short- and long-term effects using the same set of outcome measures. The following three stutter interventions were included:

1. The Comprehensive Stuttering Programme (CSP); a group therapy developed in Canada at the Institute for Stuttering Treatment and Research (ISTAR), which was introduced in the Netherlands by the HAN University of Professional Education (Hogeschool van Arnhem en Nijmegen);
2. The Doetinchem Method (DM); a group therapy given by speech and language pathologists specialised in the field of stuttering;
3. Individualised stuttering therapy according to guidelines formalised by the Association of Stuttering Treatment Centres in the Netherlands (VSN, Vereniging Stottercentra Nederland).

Two aspects of the study design need to be emphasised. First, the three stuttering programmes were not weighed against each other as such. Rather, a quasi-experimental design was adopted, with non-random assignment of stutterers to the intervention (see Frattalli, 1997). This design was chosen because of ethical considerations. In general, the choice for a specific treatment modality is largely dependent on the personality and

[2] *The research project was funded by the Dutch Health Insurance Board (CvZ, College van Zorg-verzekeringen).*

history of each client. If a stutterer were to be randomly assigned to a treatment programme that is not his primary choice, his motivation could be negatively affected and the chance of treatment success might then be significantly reduced. Secondly, the restriction to three treatment programmes is due to a limited research capacity. The treatments that were included in the study were selected on the basis of their differences in focus and approach (see subsection 2.1 for a description of the three therapy programmes).

1.3 Report layout

In section 2 the therapy programmes, the research design and methods of the present study are described in more detail. As stated above, each treatment approach was studied independently using the same protocol. Section 3 reports on our assessment outcomes per treatment programme and in Section 4 the results are summarised and placed in a more general context.

2 Method

2.1 Treatment programmes

2.1.1 The Comprehensive Stuttering Programme

The Comprehensive Stuttering Programme (CSP) is an integrated treatment programme for adolescents and adults that was developed by the Institute for Stuttering Treatment and Research (ISTAR), an affiliate of the University of Alberta in Canada. In March 2000 the CSP was introduced in the Netherlands. It was a three week intensive group-therapy program.

In the CSP, fluency-enhancing techniques are taught within a framework of prolonged speech. The techniques comprise prolongation, easy breathing, appropriate phrasing, easy onset, soft contacts, and continues airflow/blending and all exercises aim at the reduction of core stuttering and learned struggle behaviours. During the therapy the prolongation rates systematically increase from approximately 40 syllables per minute to a near normal rate of 190 syllables per minute. Clients learn to manage residual stuttering through tension modification and traditional stuttering modification techniques. Cognitive-behavioural strategies are employed to help clients (a) achieve improved communication, social skills and confidence, (b) develop positive attitudes toward communication and openness about stuttering and fluency techniques, (c) develop the ability to manage fear and anxiety, deal with negative listener reactions, and reduce avoidances, and (d) manage regression and recognise relapses. Self-management strategies are integral to all phases of the treatment programme and include problem solving, self-monitoring, and self-evaluation techniques.

There are three treatment phases: acquisition of fluency skills and cognitive-behavioural strategies, transfer of skills into non-clinical environments, and skill maintenance in the months and years following therapy. Preparation for maintenance already begins during the acquisition phase but is more emphasised in the transfer phase. Here, clients are encouraged to (1) carry out practical home assignments training fluency skills and transfer activities, (2) join or form a self-help group, and (3) seek follow-up treatment if needed (see for a full description of the CSP Kully & Langevin, 1999).

2.1.2 The 'Doetinchem Method'

The 'Doetinchem Method' (DM) is an intensive group therapy that focuses on the social perspective of stuttering, *i.e.* the social context of the stutterer is the starting-point for the therapy. The programme was developed by two Dutch speech and language therapists in 1965 and particularly targets the emotional and cognitive components of stuttering, but also aims to improve speech fluency. The programme is administered in five to six session clusters where each cluster extends over four consecutive days. Although the DM deals with stuttering from a broad perspective, it is mainly directed at two aspects of stuttering: first, the reduction of negative factors that maintain the stuttering problem and second, the enhancement of speech fluency. Based on behavioural treatment principles the client learns to describe, practise and use specific speech skills. The integration of thoughts and emotions is encouraged allowing for the clients' individual motivation and individual differences. Clients learn to reduce postural tension and tension in the respiratory muscles during speech. Stutter-free speech is trained, although improvement of communication as a whole is the main goal of the therapy. Negative emotions associated with stuttering are reduced thus boosting the client's self-esteem (Janssen, 1973; Janssen, 1994; Schoenaker & Schoenaker, 1975; van Alphen & Hoogerwerf, 1984).

Since mid-2000 the DM method as such is no longer applied. However, the DM programme has been incorporated in the so-called 'Integraal Zorg Traject' (IZT; integrated care programme). The IZT was developed by members of the DM and VSN (see next section) in 1998. Individual and group therapies were embedded within one stuttering therapy programme in order to better fulfil the individual needs of stutterers. The two DM samples that are included in the present study are the last groups that received the DM programme in its original form.

2.1.3 The VSN individual programme

Around 1995, several Dutch stuttering treatment centres joined the 'Vereniging Stottercentra Nederland' (VSN), the Association of Stuttering Therapy Centres in the Netherlands. The VSN aimed at a therapy that could be adjusted to the individual needs of each client. The association developed a programme that was firmly founded on the views of Van Riper (1973): a client-based intervention weighing the individual speech-motor components as well as the emotional and cognitive factors involved. According to the VSN stuttering is caused by a neuromuscular timing problem and because of the different (emotional and cognitive) factors that can be involved in stuttering, it needs to be approached on an individual basis, followed by group therapy if needed (de Geus, 2000, de Geus, & Putker, 2003).

In their view, stuttering originates from a hereditary predisposition to stuttering. Children with such a predisposition have a lack of accuracy in the timing of their speech production processes. When more demands on the speech production system are made (because of language developments) the stuttering problem might become manifest. This can be facilitated by three factors: (1) tension (*e.g.* resulting from high demands of the environment); (2) speaking rate (the faster the speaking rate, the more difficult the timing of the speech production processes becomes) and (3) other problems that hinder speech production (*e.g.* hearing or language problems). The three most frequently observed reactions to these timing problems are: (a) Fight (a forced attempt to prevent stuttering); (b) Flight (avoidance of words, sounds, situations and persons) and (c) Freeze (waiting, learned helplessness). The goal in therapy is to reduce these learned reactions to the core (timing) problem and thus improve communication and diminish the incidence of the dysfluencies.

The VSN individual programme is formulated according to the outcome of a diagnostic procedure prior to therapy. The treatment duration is determined by the client's individual characteristics, such as motivation, discipline, and internal locus of control (*i.e.* the extent to which the client believes that his therapy efforts will contribute towards future improvements).

2.2 *Treatment duration*

Table 2.1 shows the total number of clients per standard treatment protocol, the number of clients participating in the present study, as well as the number of clinicians per client and client-clinician contact hours for both the standard protocols and the study samples for each of the three programmes.

The CSP consisted of three consecutive treatment weeks, with a mean clinician-client ratio of 1:3.6 and an average of 25.2 contact hours per client. Since it was the first CSP clinic to be held in the Netherlands the treatment programme was organised as a 'double clinic', including 25 participants instead of the usual 12.

Table 2.1. Number of clients per treatment, clients in study, therapy duration, number of clinicians per client and client-clinician contact hours for the standard protocols and the effect study samples for all three programmes.

	CSP		*DM*		*VSN*	
	Protocol	*Effect study*	*Protocol*	*Effect study*	*Protocol*	*Effect study*
Number of clients per treatment	12	25	16	25	-	1
Clients in study	-	25	-	15	-	25**
Treatment time (in hours)	90	90	191.5	179	Variable	Variable
Clinician-client ratio	1:3.2	1:3.6*	1:8	1:6.3	1:1	1:1
Client-clinician contact hours	28.1	25.2	23.9	28.6	-	20.8(mean)

* *The therapists were 3 senior clinicians and 8 graduate students (student clinicians were counted as 0.5 senior clinicians).*

** *There were 10 dropouts during treatment, leaving 15 cases post-treatment in this study.*

The DM programme comprised five clusters of 4 days each, all administered within six months. There were two groups consisting of 10 and 15 clients. Each group had two clinicians, resulting in an average clinician-client ratio of 1:6.3. The mean number of contact hours per participant was 28.6 hours. The stutterers that participated in the two separate DM clinics (25 in total) were asked to join the effect study but only 15 clients agreed to participate.

The therapy according to the VSN is individual-based and therefore the average treatment time of the VSN programme varied across clients, ranging from 7.5 to 42 therapy hours. Sessions were mostly held once or twice a week during several months or years. The clinician-client ratio was 1:1 and there was an average of 20.8 contact hours per client.

2.3 Therapy content inventory

In order to obtain more detailed information about the actual content of the treatments administered, a so-called therapy card[3] was designed. On this therapy card the clinicians filled in (per treatment session) the actual time that was spent on the main treatment goals and the time dedicated to the strategies or skill-training exercises that were offered to achieve these goals (*e.g.* prolongation, smooth blending, confidence, social skills). The latter aspects were divided in treatment interventions targeting motor control and those directed at emotions and cognitions (see Appendix A). The average time spent on these core aspects is provided for the three programmes separately in Table 2.2. In line with the original therapy goals of the three programmes as described in the subsections of 2.1, it can be seen that the CSP had stronger focus on speech-motor control issues, whereas DM and VSN spent relatively more time on modifying emotional and cognitive aspects.

Table 2.2. Proportion of treatment time[4] spent on speech behaviour (speech-motor control) and emotions and cognitions per programme.

	CSP	*DM*	*VSN*
Speech-motor control	73.3%	36.1%	32.9%
Emotion/cognitions	26.7%	64.9%	67.1%

2.4 Participants

A total of sixty-five adult speakers who stuttered took part in the present study; there were 43 male and 22 female stutterers with a mean age of 26.5 years (age range: 17-53 years). Twenty-six[5] clients were included in the CSP group (17 men and 9 women; mean age 29.6 years, age range 17-53), 15 participants joined the DM programme (10 men and 5 women; mean age 22.3 years, ranging from 17 to 36 years) and 25 subjects participated in the VSN group (16 male and 9 female clients; mean age 27.8 years, age range 18-49). It should be noted that the sample size for the DM programme was limited due to the fact that the DM programme is no longer in use in the form described above. The two samples that were involved in this study had been part of the last two groups receiving the original DM in the Netherlands.

[3] *The therapy card was jointly designed by the researchers/clinicians involved in the present effect study and the members of the so-called 'sounding-board group' (klankbordgroep), which body consisted of representatives of the clinicians delivering the three therapy programmes and two representatives of Demosthenes, the Dutch patients' association of stutterers. In annual meetings the researchers informed the group about the state of affairs of the project.*

[4] *The percentages shown represent an estimate of the time during which the particular component was the primary focus.*

[5] *One client took part in the pretreatment assessment but never started the programme. Therefore, the actual number of participants included in the present analyses was 25.*

Because of ethical considerations the participants could not be randomly assigned to one of three treatment conditions (see section 1.2 for justification), which is why all participants were allowed to select their own programme. As a result, the educational level proved not to be equally distributed across the participants of the three programmes (see Table 2.3). Particularly in the CSP group there was a relatively high percentage of highly educated clients (68% compared to 20% for the DM and 40% for the VSN groups, when adding the percentages of pre university education and higher vocational eduction/university). This disproportion may be explained by a possible relationship between educational level and the choice for a specific type of treatment. However, such an association is highly speculative and will therefore not further be discussed.

The inclusion criteria for the present effect study were as follows:

- A reported onset of stuttering before the age of six;
- No reported problems in motor development;
- No reported concurrent problems in speech and/or language development;
- No reported use of medication that could influence respiration, phonation or articulation;
- No reported psychiatric problems;
- No reported hearing problems.

Before the start of treatment the clients were extensively informed about the study and invited to participate. The participants who were prepared to join the study subsequently signed an informed consent prior to the first therapy session.

Table 2.3. Number (and proportion) of stutterers per educational level for each treatment programme.

	Lower vocational education	*Intermediate vocational education*	*Pre university education*	*Higher vocational education and university*	*Total*
CSP	1 *(4%)*	7 *(28%)*	3 *(12%)*	14 *(56%)*	25 *(100%)*
DM	4 *(26.7%)*	8 *(53.3%)*	1 *(6.7%)*	2 *(13.3%)*	15 *(100%)*
VSN	5 *(20%)*	10 *(40 %)*	- *(0%)*	10 *(40%)*	25 *(100%)*

2.5 Assessment schedules

2.5.1 Short- and long-term treatment effect measures

It is generally known that there are treatment programmes that yield spectacular improvements in the speech of stutterers during or immediately after therapy, in other words the short-term effects for these interventions are quite favourable. However, it is also known that many participants show regression or even experience a complete relapse one or two years later. To gain a deeper insight into these long-term treatment effects, Bloodstein (1987) recommends including 18- or 24-months follow-up measurements in efficacy studies. Therefore, in the present study, participants were tested immediately before the start of the therapy (pretreatment assessment), immediately after conclusion of the programme (post-treatment assessment), one year following termination of treatment (first follow-up assessment, F1), and finally two years after therapy conclusion (F2).

2.5.2 Non-treatment-related variability

All speakers show a normal ‘day to day’ variation in their speech behaviour, and so do stuttering speakers (Alfonso & Van Lieshout, 1997). In our study, this variation in individual speech characteristics was measured to allow a better interpretation of the differences between pre- and post-treatment data. The non-therapy related variability was tested in 12 clients (5 women and 8 men with a mean age of 25.6 years, ages ranging between 17 and 49 years) on two different occasions before the start of the therapy (with approximately one month between the two pretreatment measurements). The two datasets were compared with each other to establish variation in speech behaviours across time. To this end, speech samples (CD-recordings) of a monologue and reading task were judged using the perceptual rating instrument developed by Franken, Boves, Peters, and Webster (1995; see section 2.6.1.2 for a description of the rating tool) by ten naïve listeners (5 women and 5 men, mean age 21.9 years, age-range 19 to 30 years). The results revealed that there were no significant differences between the two pretreatment assessments.

2.6 Variables

2.6.1 Speech behaviour

2.6.1.1 Stutter severity

Syllables per minute (SPM) and percentage of stuttered syllables (%SS)

Speech rate, as expressed by the number of syllables produced per minute (SPM), was obtained by counting the number of syllables uttered in four 2-minute speech samples (an interview, reading task, monologue, and a telephone conversation) recorded at the pre- and post-treatment, F1 and F2 assessments. All syllable counts were made using electronic button-press event recorders (see Boberg and Kully, 1985, 1994). Overt stuttering severity was assessed as a percentage of syllables stuttered using the following formula: total syllables stuttered/total syllables x 100% = SS%. To this end, for each participant and for each assessment the number of stutters occurring in the two minutes of each speech sample was counted.

SPM and %SS reliability

Three trained raters (for a description of the training session for the raters, see Boberg and Kully, 1994) counted the number of dysfluencies in the three-minute speech samples of the four speech tasks for all participants. Individual data were randomly assigned to the raters with stratification for therapy, *i.e.* the speech samples were randomly assigned to the raters but one rater evaluated all four measurements (pre, post, F1 and F2) of each selected individual participant. Inter- and intra-reliability was calculated with the intra-class correlation coefficient (see Shrout and Fleiss, 1979). For inter-rater reliability, 159 (16.70%) speech samples from a total of 952 samples (238 sessions x 4 speech tasks; also see Table 3.1) were recounted by all three raters. For intra-rater reliability, 112 (11.76%) samples were recounted by the same three raters approximately 6 months after the first assessment. Table 2.4 shows the inter-rater reliability and intra-rater reliability scores. The high scores correspond to the score levels Langevin and Boberg (1993) and Boberg and Kully (1994) obtained.

Table 2.4. Intra-rater reliability and inter-rater reliability per therapy programme for syllables per minute (SPM) and the percentage of stuttered syllables (%SS).

	Intra-rater reliability			*Inter-rater reliability*		
	Rater 1	*Rater 2*	*Rater 3*	*CSP*	*DM*	*VSN*
SPM	0.953	0.992	0.969	0.957	0.969	0.988
%SS	0.988	0.994	0.985	0.979	0.958	0.995

Self-evaluation: the speaker's own judgment

Participants were requested to judge their own speech on a speech satisfaction scale. A ten-point rating scale was used, ranging from 1 to 10, with 1 as the worst possible judgment and 10 as the best possible judgment. This is in line with the grading system employed in all educational institutes in the Netherlands, where five and below five represent insufficient and six and above six sufficient performance (comparable to A-F grades in the English and North-American grading system).

2.6.1.2 Speech quality

The perceptual rating instrument developed by Franken et al. (1995) and Franken, Boves, and Peters (1997) was used in the present study to examine changes in speech quality. The instrument consists of 14 speech characteristics that are evaluated with respect to phonation, articulation, loudness and naturalness on a seven-point scale. The speech aspects are rated by means of bipolar scales that are defined by contrastive terms labelling extremes (analogous to the Semantic Differential Scale composed by Osgood, Succi, and Tannenbaum, 1957). The 14 speech characteristics that are assessed are: high pitch-low pitch; quick-slow; slovenly-polished; expressive-flat; shrill-deep; soft-loud; melodious-monotonous; tense-relaxed; weak accentuation-strong accentuation; unpleasant-pleasant; slurred-precise; halting-fluent; weak-powerful; natural-unnatural. For our evaluations, we selected 45 seconds from the speech samples recorded during the monologue task the participants had carried out at each assessment (pre, post, F1 and F2). We subsequently chose randomly 39 speech samples, *i.e.* 13 from each therapy programme. These restrictions were prompted by practical reasons (containment of the duration of the rating session to prevent the raters from becoming overtaxed). Thus the total number of speech samples was 156 (3 therapy programmes x 13 stutterers x 4 assessments). Using the perceptual rating instrument, forty-two untrained female listeners with a mean age of 20.2 years (age range 17-26 years) evaluated the speech samples. All raters were first year logopaedics students without prior experience in assessing or treating stuttering and therefore considered as essentially naïve with respect to the formal and technical aspects and terminology of evaluating speech samples in this population. The speech samples were randomised with the restriction that two samples of the same speaker had to be separated by at least three samples of other speakers. All speech samples (including four practice samples) were presented on two cds. The ratings were organised in the form of a classroom session with a total duration of 3 hours, including the instruction and three breaks. The average ratings were used in the analysis (for justification, see Franken et al., 1995). The reliability of this average was very high (each scale displayed an intra-class correlation coefficient >0.95).

In order to understand the dimensional structure of the fourteen scales, a Factor Analysis (Principal axis factoring with Varimax Rotation) was carried out on the F2 dataset, using SPSS 11.01. This resulted in a three-factor solution, which is shown in Table 2.5. The first factor can be defined as *Articulation Quality* with high loadings on the scales Unpleasant-Pleasant, Halting-Fluent, Unnatural-Natural, Tense-Relaxed, Slurred-Precise and Slovenly-Polished. The second factor is taken to reflect *Voice Dynamics* with high loadings on the scales High-Low Pitch, Shrill-Deep, Melodious-Monotonous, Expressive-Flat and Quick-Slow.
The third factor is classified as the *Speech Power* factor with high loadings on the scales Soft-Loud, Weak-Powerful, and Weak-Strong Accentuation. The factor scores were saved using the Bartlett method, *i.e.* an estimation of the factor score coefficients.

Table 2.5 Factor analysis of the averaged scores of the 42 raters on the 14 ratings scales.

Rating Scales	*Articulation Quality*	*Voice Dynamics*	*Speech Power*
Unpleasant-Pleasant	**0.95**	-0.19	0.16
Halting-Fluent	**0.94**	-0.02	0.17
Unnatural-Natural	**0.93**	-0.17	0.25
Tense-Relaxed	**0.93**	0.04	0.07
Slurred-Precise	**0.83**	-0.32	0.16
Slovenly-Polished	**0.59**	-0.32	0.14
High Pitch-Low Pitch	-0.09	**0.97**	-0.02
Shrill-Deep	-0.10	**0.88**	-0.22
Melodious-Monotonous	-0.49	**0.63**	-0.54
Expressive-Flat	-0.45	**0.61**	-0.58
Quick-Slow	-0.10	**0.41**	-0.39
Soft-Loud	-0.05	-0.10	**0.91**
Weak-Powerful	0.44	-0.12	**0.89**
Weak Accentuation-Strong Accentuation	0.45	-0.44	**0.65**

The scores produced have a mean of 0. The sum of squares of the unique factors over the range of variables is minimised. Next, these factor scores were used to evaluate the effect of treatment. It should be noted that a high score on the Articulation Quality factor signifies good (pleasant, fluent, natural, precise and polished) articulation. A low score on the Voice Dynamics factor denotes dynamic speech (high pitch, shrill, melodious, expressive and quick) and a high score on the Speech Power factor reflects powerful speech (loud, powerful and strong accentuation).

2.6.1.3 Speech-motor control

To study speech-motor control, the speech-capacity subtests from the Nijmegen Speech Motor Test (NSMT) were used (see Peters, 1999; Van Lieshout, Peters, & Bakker, 1997 for an introduction to the NSMT and its subtests). In the NSMT, specific speech tasks are assessed with regard to basic speech-motor skills. For example, the maximum repetition rate is evaluated in a diadochokinesis task (DDK), speech-motor complexity is measured by varying word length, cluster complexity (see Huinck, Wouters, Hulstijn, & Peters, 2001; Diepstra, Huinck, Hulstijn, & Peters, 2001), cluster location, tongue twisters (see Huinck, Van Lieshout, Peters, & Hulstijn, 2004) and sentence accent (or stress). Linguistic influences are assessed by comparing words with non-words (see Huinck et al., 2001) and respiration behaviour is tested in a task in which response time is varied such that respiration needs to be fully controlled. In most of these tasks, reaction times are measured in a delayed and/or immediate reaction-time paradigm (see also Van Lieshout, Hulstijn, & Peters, 1996a, 1996b). Changes in speech-motor characteristics are measured through simultaneous recordings of peripheral physiological processes: respiration is determined with the Respitrace™, phonation with the Fourcin® Electrolaryngograph (EGG) and articulation is measured by surface electromyographic (sEMG) recordings of the activity of the m. orbicularis oris superior and inferior.

For the present report, the DDK task was used as a measure of overall speech-motor skill and also to test the maximum repetition rate of the speech-production system (see Ackermann, Hertrich, and Hehr, 1995; Ackermann, Konczak, & Hertrich, 1997). Participants had to repeat three different syllable sequences, /pə/, /təkə/, and /pətəkə/, as quickly as possible. The number of correctly produced syllables for the three sequences was counted during 5, 7 and 9 seconds, respectively. Each sequence had to be repeated twice and the mean number of correctly produced syllables was taken as the dependent variable. We will report on the data analyses of the remaining tasks of the NSMT in a future publication.

2.6.2 Self-report questionnaires

In each session (pre, post, F1 and F2), the participants completed several self-assessment questionnaires in which they evaluated changes in their feelings and emotions that are assumed to be associated with stuttering.

The first of the 6 self-report questionnaires that were used was the *Perceptions of Stuttering Inventory (PSI;* Woolf, 1967). This test comprises 60 items that represent parameters of struggle (S), avoidance (A) and expectancy (E). The PSI examines the stutterer's perception of the presence of these factors in his communication. Woolf suggests that the PSI can be used to help the stutterer view his problem more objectively, to develop treatment goals and to assess progress. Levels of severity are distinguished by the following division: scores below 7 are termed as mild, those between 8 and 11 are designated as moderate, 12 to 15 as moderate-to-severe and scores from 16 to 20 are categorised as severe. Thus, higher scores on this questionnaire indicate a more negative perception of stuttering.

Our second self-report tool was the *Speech Situation Checklist (SSC;* Brutten, 1975; Brutten & Janssen, 1981). This questionnaire measures the self-reported emotional reaction (ER) and the degree of speech distortion (DS) in 54 concrete speech situations that have been associated with different levels of negative emotion and fluency failure. Respondents indicate ER and DS on a five-point scale. The higher the score, the higher the ER or DS on the given situation. The scores are compared with the norm scores of stuttering speakers (mean ER=129.3, SD=30.5; mean DS=129, SD=29.3) and those of non-stuttering speakers (Brutten & Janssen, 1981).

Also, the *S44* (Erickson, 1969) and the *Modified Erickson Scale (S24;* Andrews & Cutler, 1974) were administered. These scales measure attitudes towards communication ability. Higher scores reflect more negative attitudes. The norm score on the S44 is 28.4 (SD=8.3) and for the S24 this is 19.2 (SD=4.2).

Furthermore, Lanyon's *Stuttering Severity Scale (SS Scale;* 1967) was completed. This inventory of stuttering-related behaviours and attitudes comprises 64 speech situations that assess the severity of stuttering. The stutterer is asked to indicate whether the true/false statements (example: "When I talk, I often become short of breath") are applicable to him/her. The scores are compared with a norm score of stuttering speakers (40.6; SD=11.9).

In addition, we used the *Inventory of Interpersonal Situations (IIS)* developed by van Dam-Baggen and Kraaimaat (1987). This questionnaire measures two components of social anxiety: the extent to which emotional tension or discomfort is perceived in social situations and the frequency with which social responses are executed. Both parts employ the same 35 items to elicit responses to social situations. The 35 statements are grouped in five subscales: 'giving criticism', 'expressing opinion', 'giving a compliment', 'initiating contact', and 'positive self-statements'. Overall, stuttering speakers display significantly higher levels of emotional tension or discomfort in social situations than non-stuttering speakers (Kraaimaat, Vanryckeghem & van Dam-Baggen, 2002).

Our sixth questionnaire comprised the *Achievement Motivation Test (AMT;* Hermans, 1968, 1970). This test is designed to investigate the respondent's characteristics that are related to achievement motivation in different situations and was not specifically developed to assess stuttering speakers. The AMT consists of three subscales: (1) the Achievement scale (A) which measures the respondent's personality with respect to achievement, (2) the negative failure anxiety scale (F-) which assesses anxiety that causes dysfunction, and (3) the positive failure anxiety scale (F+) which gauges the anxiety that is beneficial to the respondent's functioning.

The seventh and final self-assessment tool we employed was the *Speech Performance Questionnaire (SPQ)* The SPQ (Perkins, 1981) was designed to assess the clients' perceptions of their post-treatment performance. Researchers of the Canadian ISTAR group have developed an adapted version of the SPQ (see Langevin and Boberg, 1993). Our participants completed this latter version during the F1 and F2 assessments. With the SPQ stutterers can evaluate various aspects of their speech performance and indicate levels of satisfaction during the follow-up periods. Table 2.6 categorises the questionnaires that were used in this study.

Table 2.6. Overview of the focus of the different questionnaires.

	Brutten SSC	*Erickson S-scale*	*Lanyon SS-scale*	*Woolf PSI*	*IIS*	*Herman AMT*	*Perkins SPQ*
Stutter severity	X	X	X	X			
Speech anxiety	X			X			
Attitude		X					
Social anxiety					X		
Fear of failure						X	
Satisfaction							X

2.7 Statistical analyses

For each therapy programme the data were analysed using a repeated measures design (SPSS 10, GLM repeated measures) with therapy session (Pre-, Post-treatment, F1 and F2) as within-subject factor. To study the differences between pre- and post-treatment, pretreatment and F1, pretreatment and F2, post-treatment and F1, and between F1 and F2, pair-wise comparisons (Bonferroni[6] corrected for multiple comparisons) were carried out.

To analyse the factor scores concerning perceptual evaluation, again, a repeated measures design was used. When there was a significant session effect, simple contrasts between pre- and post-treatment, F1 and F2 are reported. The significance level was set at 0.05.

The results of our analyses are reported in Section 3. Because the three therapy programmes were not compared with each other, the results are presented for each treatment separately.

3. Results

3.1 Drop-outs

Although in patient-related research it is a well-known problem that a proportion of clients abandon their treatment prematurely, there is little literature on the frequency and nature of drop-out behaviour for stuttering therapy. It is generally assumed that most of the stuttering adults who have decided to attend a stuttering programme are highly motivated and thus complete the whole treatment as prescribed by the therapists. Nevertheless, even in this highly motivated population there are clients that quit during the course of their therapy. The grounds for dropping out are very diverse. Sometimes there are practical reasons (*e.g.* prolonged commute due to a move or lack of time due to a new job) but it may also result from a waning motivation or a lack of

[6] *Because of the explorative nature of this study we did not correct for the different measures that were used.*

progress. It is obvious that the longer the duration of the therapy, the higher the chance of dropping out becomes.

Table 3.1 shows the number and percentage of drop-outs in relation to the number of participants in the pretreatment assessment for each subsequent assessment and for each treatment programme. Eleven (16.7%) of the original 66 clients who were assessed prior to treatment never participated in the post-treatment measurement. One year later (F1) this number had increased to 13 (19.7%), and at the two-year follow-up (F2) there were another 14 clients that had ended their contribution to the study. Thus, the total drop-out rate was 21.2%.

The highest percentage of drop-outs was found in the VSN group. This may be related to the individual nature of the programme and thus its potentially prolonged time-span (sometimes covering even more than two years), which increases the probability that the client's living or working situation changes during this period. It is also possible that group dynamics played a role in the group therapies in that it kept the individual participants motivated to complete the programme. However, these interpretations are tentative and have not been verified systematically.

Table 3.1. Number of participants (and percentage of drop-outs in relation to pretreatment levels) for the Pre-treatment (Pre), Post-treatment (post), Follow-up 1 (F1) and Follow-up 2 (F2) assessments for each treatment programme. The total reflects the number of individual observations for the entire study.

	Extra pre measurement	*Pre*	*Post*	*F1*	*F2*	*Total*
DM	0	15	15 *(0%)*	14*(6.7%)*	14*(6.7%)*	*58*
VSN	0	25	15 *(40%)*	14 *(44%)*	13 *(48%)*	*67*
CSP	13	26[5]	25*(3.8%)*	25*(3.8%)*	25*(3.8%)*	*113*
Total	*13*	*66*	*55(16.7%)*	*53(19.7%)*	*52(21.2%)*	*238*

3.2 Comprehensive Stuttering Programme

3.2.1 Speech behaviour

3.2.1.1 Stutter severity

Number of syllables per minute (SPM)

In Table 3.2 the mean SPM and mean differences relative to the pretreatment measurements are shown. Overall, the SPM increased significantly after therapy (p<0.01). Although there was regression in F1 (15%) and F2 (10.1%), the difference between pretreatment and F1 (p<0.05) and F2 (p<0.01) was maintained.

Table 3.2. Mean number of syllables per mintute (SPM), percentage stuttered syllables (%SS) and differences between pre- and post-treatment, between pretreatment and Follow-up 1 and between pretreatment and Follow-up 2, averaged over the four speech tasks (interview, monologue, reading, telephone conversation) for the CSP group.

Measurement session	*Mean and differences relative to pre-treatment (*p<0.05; **p<0.01)*			
	SPM		*%SS*	
	Mean	*Differences relative to Pre-treatment*	*Mean*	*Differences relative to pretreatment*
Pretreament	126.0		13.2	
Post-treatment	166.7	-40.7 **	1.2	12.0 **
Follow-up 1	147.8	-21.8 *	8.0	5.2 **
Follow-up 2	154.0	-28.0 **	6.9	6.3 **

Figure 3.1 shows the SPM[7] values for the CSP group. The bars show the mean SPM per session for each speech task. There are significant differences between the pre- and post-treatment values for the tasks involving the interview (p<0.05), the monologue (p<0.01) and the telephone conversation (p<0.01), but not for the reading task. There was a significant gain between pretreatment and F1 performance for the monologue (p<0.05) and the telephone conversation (p<0.01). The difference between pretreatment and F2 values was significant for the interview (p<0.05), the monologue (p<0.01) and the telephone conversation (p<0.01). Hence, reading was the only task that failed to reveal any significant improvements after therapy.

Percentage of stuttered syllables (%SS)

Overall, the %SS had decreased significantly (p<0.01) after therapy (Table 3.2) and again, in spite of the regression that occurred in the long term (51.3% in F1 and 43% in F2), an important proportion of the gain achieved between pretreatment and F1 and between pretreatment and F2 was maintained (p<0.01). In Table 3.2 the mean %SS and the mean differences relative to the pretreatment levels are reported. Figure 3.2 shows the mean and standard error (SE) %SS for each of the four assessments for the CSP group. There are significant differences between the pre- and post-treatment

[7] *There is no literature available on the norm SPM for non-stuttering Dutch speakers. Janse (2004) calculated a normal articulation rate of 6.1 syllables/s in a reading task. However, because of the different method that was used in the present study the two speech rates are not compatible.*

values on all four speech tasks ($p<0.01$). The gains between pretreatment and F1 and between pretreatment and F2 were also significant ($p<0.01$).

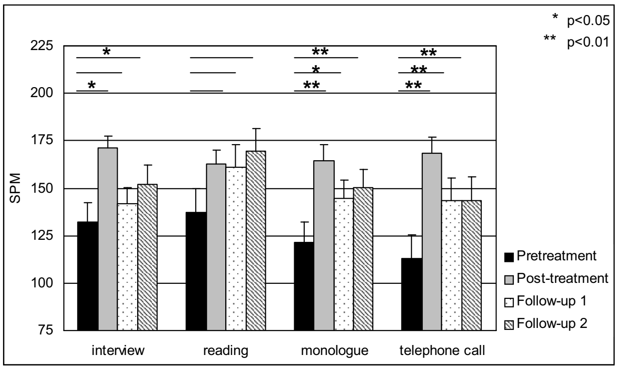


Figure 3.1. Mean number (and SE) of syllables per minute (SPM) for the interview, reading task, monologue and telephone call for the CSP group for all four assessments.

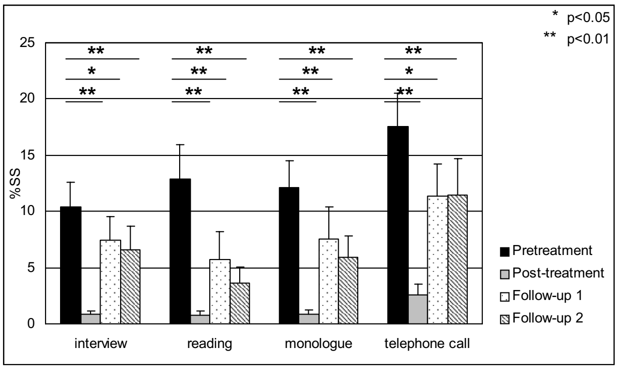


Figure 3.2. Mean percentage (and SE) of stuttered syllables (%SS) for the CSP group for the interview, reading task, monologue and telephone call for each measurement.

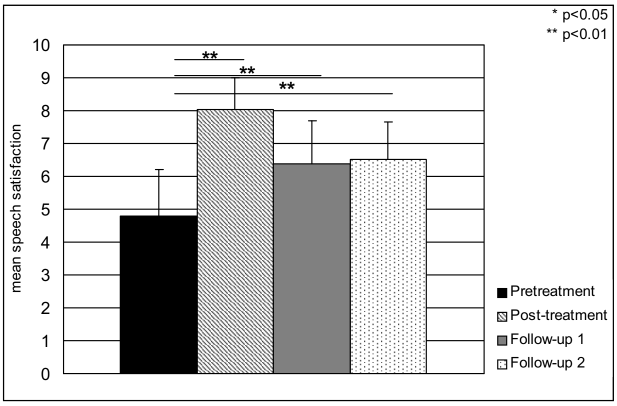


Figure 3.3. Mean scores (and SDs) on the speech satisfaction scale for the CSP group for each measurement.

Self-evaluation
Figure 3.3 shows the mean scores on the satisfaction rating scale (from 1 to 10) for the CSP group. At the pretreatment assessment, the mean rating of the participants was 4.79 (SD=1.43). At the post-treatment measurement this rating had increased significantly ($p<0.01$) to 8.04 (SD=0.96). At the one- and two-year follow-ups the scores indicated regression (6.39; SD=1.29 and 6.51; SD=1.13, respectively) but, relative to the pretreatment levels, both gains remained significant ($p<0.01$).

3.2.1.2 Speech quality
Figures 3.4a, b and c show the mean scores of the CSP group for each of the three factors. On the factor Voice Dynamics (Figure 3.4a), the score had increased at post-treatment assessment and decreased again in F1 and F2. As mentioned above, a lower score on this factor signifies more dynamic speech. There was no main effect of session (pre, post, F1 and F2) and thus no treatment effect. Figure 3.4b reflects the mean score for the factor Articulation Quality. The higher the score on this factor the better the speech. The main effect session (pre, post, F1 and F2) was significant ($F(3,36)=4.974$; $p<0.001$). At post-treatment and relative to pretreament, the factor value had improved significantly ($F(1,12)=8.354$; $p=0.014$). This effect was sustained in both F1 ($F(1,12)=9.191$; $p=0.010$) and F2 ($F(1,12)=6.634$; $p=0.024$), which indicates that the short- and long-term gains in the quality of articulation were maintained. In Figure 3.4c the mean score on the Speech Power factor is depicted. The higher the score on this factor, the more powerful the speech was rated. The analysis revealed that none of the four assessments showed a significant main effect of session.

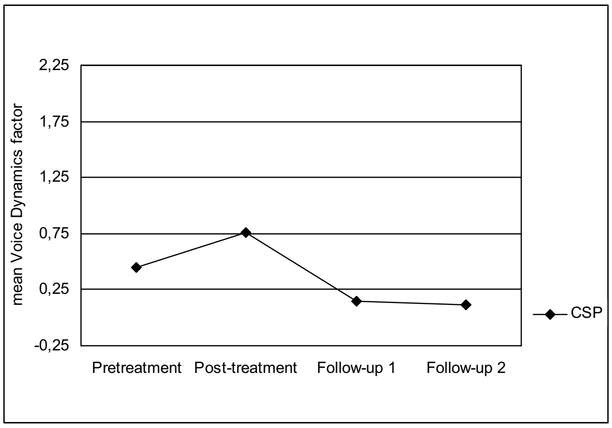


Figure 3.4a. Mean scores of the CSP group on the Voice Dynamics factor for each assessment (pre, post, F1 and F2).

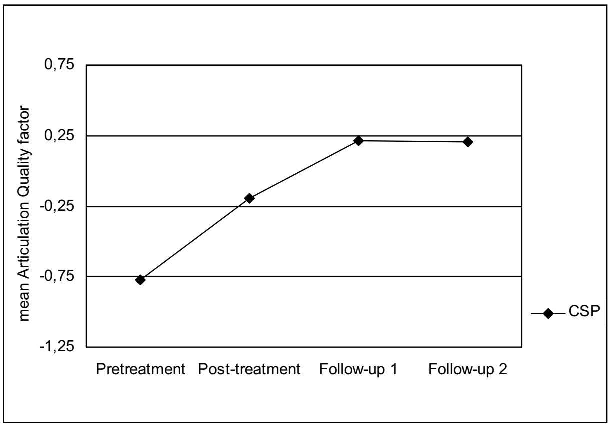


Figure 3.4b. Mean scores of the CSP group on the Articulation Quality factor for each assessment (pre, post, F1 and F2).

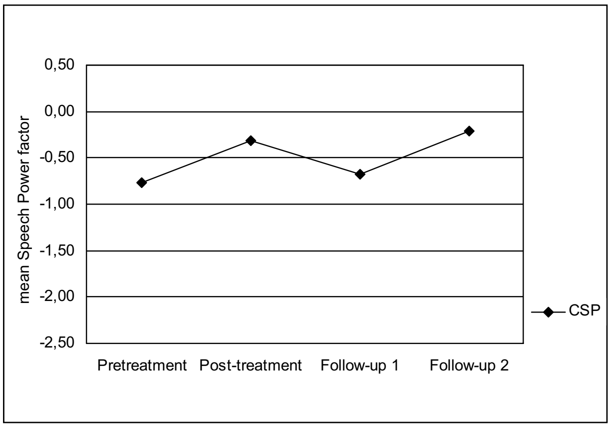


Figure 3.4c. Mean scores of the CSP group on the Speech Power factor for each assessment (pre, post, F1 and F2).

3.2.1.3 Speech-motor control

At the post-treatment assessment, the maximum repetition rate had increased significantly ($p<0.01$) on the sequence /pətəkə/. This treatment effect was maintained in both F1 ($p<0.01$) and F2 ($p<0.05$). There were no post-treatment changes on the sequences /pə/ and /təkə/ (see Table 3.3).

Table 3.3. Mean scores (number of correctly produced syllables) and differences relative to pretreatment for the CSP group per sequence in the DDK task.

	*Mean scores and mean differences relative to pre-treatment for the DDK (*p<0.05, **p<0.01)*					
	pə (SE)		*təkə (SE)*		*pətəkə (SE)*	
CSP (n=24)	*Mean*	*Mean differences*	*Mean*	*Mean differences*	*Mean*	*Mean differences*
Pretreatment	23.23		17.85		13.19	
Post-treatment	23.35	-0.12	18.65	-0.80	16.15	-2.96**
Follow-up 1	24.96	-1.73	18.75	-0.90	16.75	-3.56**
Follow-up 2	24.40	-1.17	17.17	0.68	15.77	-2.58*

3.2.2 Self-report questionnaires

The responses to the self-report questionnaires are summarised in Table 3.4. The table shows the mean differences between pre- and post-treatment, pretreatment and follow-up 1 and between pretreatment and follow-up 2 values. Significant gains are indicated with *($p<0.05$) and **($p<0.01$).

(1) Perceptions of Stuttering Inventory (PSI)

The mean score of the CSP group on the PSI was: Struggle (S)=9.91 (SE=1.14); Avoidance (A)=9.24 (SE=1.09); Expectancy (E)=9.24 (SE=0.91). Post-treatment struggle behaviour was lower ($p<0.01$). Although there was regression in the long term, the effects, relative to pretreatment, were maintained in both F1 and F2 ($p<0.05$). Post-treatment avoidance behaviour had also decreased relative to pretreatment measures ($p<0.01$). Again there was little regression in the long term and the gains were maintained in both F1 $p<0.01$) and F2 ($p<0.01$). Expectancy had also decreased at post-treatment assessment ($p<0.05$). In F1 this effect disappeared but in F2 there was again a significant difference relative to the pretreatment value ($p<0.05$).

(2) Brutten's Speech Situation Checklist (SSC)

The norm score for stutterers on the Emotional Reaction scale (ER; this is the total ER score, *i.e.* the sum of the various ER dimensions) was 129.3 (SD=30.5) and on the Distorted Speech scale (DS; this is the total DS score, *i.e.* the sum of the various DS dimensions) 129 (SD=29.3).

The CSP group showed a mean pretreatment ER of 152.7 (SD=15.4) and at post-treatment measurement this had decreased significantly ($p<0.01$) to 101.3 (SD=22.8); in F1 there was some regression (130.1; SD=20.9) but, relative to the pretreatment score, the difference remained significant ($p<0.05$). In F2 the mean score was 121.8 (SD=22.4), again a significant gain compared with the pretreatment score ($p<0.05$). Thus, despite a moderate regression after one year, the post-treatment reduction in emotional reactions to the presented speech situations was maintained in the long term.

The mean pretreatment DS was 153.1 (SD=18.5). There was a significant reduction at post-treatment assessment ($p<0.01$) to 96.3 (SD=25.4). In F1 this was 130.9 (SD=22.3) and in F2 125.6 (SD=17.2). Thus again, despite considerable regression, the long-term effects were maintained in both F1 and F2 ($p<0.05$).

(3) S24 and S44

Prior to therapy, the mean score of the CSP group was 27.14 (SE=1.59) on the S44 and 15.73 (SE=0.99) on the S24. At the post-treatment assessment these scores had decreased significantly (p<0.01) to 17.27 (SE=1.47) for the S44 and 9.46 (SE=0.93) for the S24. In F1 there was a moderate regression to a mean score of 20.46 (SE=1.53) and 11.68 (SE=0.96), respectively. In F2 this was 18.77 (SE=1.44) and 10.77 (SE=0.81), respectively. Despite the slight intermediate regression, the gains on both the S24 and the S44 were maintained (p<0.01) in F1 and F2.

(4) Lanyon's Stuttering Severity Scale

The norm score on this questionnaire was 40.6 (SD=11.9). The mean score of the CSP group was 39.48 (SE=2.28). At post-treatment this score had decreased significantly (p<0.01) to 24.52 (SE=2.12). Relative to the pretreatment values there was regression in F1 (30.67; SE=2.70) and F2 (28.90; SE=2.52) but the effects remained significant (p<0.01).

(5) Inventory of Interpersonal Situations (IIS)

On the IIS the participants had to indicate the degree of tension that they associated with the 35 situations described and also how often (frequency) these situations occurred. At the post-treatment assessment the tension scores had decreased on four of the five subscales: 'Criticizing' (p<0.01), 'Getting attention' (p<0.01), 'Appreciation' (p<0.05) and 'Initiating a conversation' (p<0.01). The subscale 'Appreciating yourself' had not decreased significantly. Treatment effects were maintained in the long term on the subscales 'Getting attention' (p<0.01) and 'Initiating a conversation' (F1:p<0.05; F2:p<0.01). On the IIS frequency scale the stutterers reported increased frequency for 'Criticising' (p<0.01), 'Getting attention' (p<0.01), 'Appreciation' (p<0.05) and 'Initiating a conversation' (p<0.01).
The gains were also maintained in the long term for the scales 'Criticising' (p<0.01), 'Getting attention' (p<0.05) and 'Initiating a conversation' (p<0.01).

(6) Hermans' Achievement Motivation Test (AMT)

The CSP group had a mean pretreatment achievement motivation (A) of 15.10 (SE=1.51), a mean negative fear of failure (F-) of 13.95 (SE=0.88) and a mean positive fear of failure (F+) of 8.24 (SE=0.89). There was no effect of therapy on the scales A and F-. In contrast, the scores for F+ were reduced at the post-treatment assessment and in F2 this decrease proved significant (p<0.05).

The Speech Performance Questionnaire (SPQ)

Table 3.5 shows for each item on the SPQ how often the particular score was reported in both F1 and F2. Sixty-three percent of the 22 clients in F1 and 74% of the 23 clients in F2 rated their current speech fluency as 'very good' or 'generally good' (item 4). In both F1 and F2 83% of the clients found the CSP 'very helpful' or 'moderately helpful' (item 12) and 100% of the clients in F1 and 86% of the clients in F2 attributed the speech improvements to the CSP programme (item 15). Seventy-five percent of the clients in F1 and 78% of the clients in F2 labelled their confidence in speech as 'much improved' or 'moderately improved' (item 18).

Table 3.4 Mean differences (and standard errors in parentheses) between pre- and post-treatment assessment, pretreatment and Follow-up 1 and pretreatment and Follow-up 2 on the self-report questionnaires for the CSP group.

Questionnaires	*Subscales*	*Mean Differences CSP (*p<0.05, **p<0.01)*		
		Pre-Post (SE)	*Pre-F1(SE)*	*Pre-F2(SE)*
Perceptions of Stuttering Inventory (PSI)	*Struggle*	6.10 (1.08)**	2.29 (1.07)*	3.29 (1.11)*
	Avoidance	6.43 (1.31)**	4.00 (1.36)**	4.67 (1.45)**
	Expectancy	3.29 (1.24)*	1.52 (1.29)	3.33 (1.16) *
Brutten's SSC	*Emotional Reaction*	1.01 (0.15)**	0.44 (0.19)*	0.61 (0.19)*
	Distorted Speech	1.11 (0.18)**	0.44 (0.17)*	0.54 (0.17)*
Erickson/ Andrews & Cutler	*S44*	9.86 (1.96)**	6.68 (1.63)**	8.36 (1.75)**
	S24	6.27 (1.17)**	4.05 (1.01)**	4.96 (0.94)**
Lanyon's Stuttering Severity Scale		14.95 (2.95)**	8.81 (2.75)**	10.57 (2.53)**
Inventory of Interpersonal Situation (IIS) Tension	*Criticising*	3.29 (1.05)**	1.76 (1.04)	1.29 (1.02)
	Getting attention	4.24 (0.94)**	2.90 (0.94)**	3.33 (0.85)**
	Appreciation	0.95 (0.45)*	0.14 (0.47)	0.38 (0.39)
	Initiating a conversation	3.19 (0.82) **	1.76 (0.68)*	2.43 (0.69)**
	Appreciating yourself	0.71 (0.69)	0.81 (0.70)	0.76 (0.58)
	Total	15.71 (4.54)**	10.14 (4.34)*	10.43 (3.43)**
Inventory of Interpersonal Situation (IIS) Frequency	*Criticising*	-3.00 (0.57)**	-2.62 (0.89)**	-2.57 (0.83)**
	Getting attention	-2.48 (0.62)**	-2.43 (0.87)*	-1.57 (0.66)*
	Appreciation	-1.19 (0.46)*	-0.86 (0.57)	-0.52 (0.64)
	Initiating a conversation	-2.90 (0.57)**	-2.10 (0.58)**	-2.76 (0.40)**
	Appreciating yourself	-1.43 (0.69)	-0.67 (0.62)	-0.90 (0.63)
	Total	-14.29(2.45)**	-11.67 (3.26)**	-10.67(1.95)**
Hermans' AMT	*Achievement Motivation (A)*	-0.57 (0.71)	-0.52 (0.83)	-0.19 (0.95)
	Neg. Failure Anxiety (F-)	0.43 (0.71)	0.57 (0.51)	1.05 (0.71)
	Pos. Failure Anxiety (F+)	-0.95 (0.56)	-1.05 (0.92)	-1.71 (0.76)*

Table 3.5. Number of responses per scale (1, 2, 3 and 4) for each item of the SPQ at F1 and F2 for the CSP sample.

CSP Responses to the SPQ (F1: n=24; F2: n=23)	*Number of responses* **F1** *and F2*			
	1	***2***	***3***	***4***
1. Satisfaction with speech before therapy	**0**	**4**	**15**	**5**
(1. very satisfied, 2. generally satisfied, 3. generally dissatisfied, 4. very dissatisfied)	*0*	*4*	*13*	*6*
2. Satisfaction with speech immediately after therapy	**16**	**7**	**0**	**1**
(1. very satisfied, 2. generally satisfied, 3. generally dissatisfied, 4. very dissatisfied)	*17*	*5*	*0*	*1*
3. Current rating of speech satisfaction	**1**	**13**	**8**	**2**
(1. very satisfied, 2. generally satisfied, 3. generally dissatisfied, 4. very dissatisfied)	*0*	*16*	*4*	*3*
4. Current rating of speech fluency	**1**	**13**	**7**	**1**
(1. very good, 2. generally good, 3. generally dissatisfied, 4. very dissatisfied)	*0*	*17*	*5*	*1*
5. Now have necessary skills to control speech	**3**	**11**	**9**	**1**
(1. almost always, 2. most of the time, 3. some of the time, 4. seldom)	*2*	*11*	*7*	*3*
6. Now have necessary skills to sound fluent	**4**	**11**	**8**	**1**
(1. almost always, 2. most of the time, 3. some of the time, 4. seldom)	*3*	*11*	*7*	*2*
7. Now have necessary skills to sound normal	**5**	**10**	**5**	**4**
(1. almost always, 2. most of the time, 3. some of the time, 4. seldom)	*3*	*12*	*6*	*2*
8. I use my speech controls	**0**	**3**	**15**	**6**
(1. almost always, 2. most of the time, 3. some of the time, 4. seldom)	*0*	*4*	*15*	*4*
9. Now able to speak normally without thinking about controls	**2**	**5**	**8**	**9**
(1. almost always, 2. most of the time, 3. some of the time, 4. seldom)	*3*	*7*	*10*	*3*
10. Now feel like a normal speaker	**1**	**5**	**11**	**7**
(1. almost always, 2. most of the time, 3. some of the time, 4. seldom)	*1*	*11*	*8*	*3*
11. As a result of therapy my speech fluency is	**5**	**8**	**7**	**4**
(1. much improved, 2. moderately improved, 3. slightly improved, 4. not improved)	*8*	*4*	*8*	*3*
12. I found the CSP therapy programme to be	**11**	**9**	**3**	**1**
(1. very helpful, 2. moderately helpful, 3. slightly helpful, 4. not helpful)	*8*	*11*	*4*	*0*
13. Prefer stuttering over controlled speech	**0**	**3**	**11**	**9**
(1. all of the time, 2. most of the time, 3. some of the time, 4. seldom)	*0*	*4*	*8*	*11*
14. Currently consider myself a stutterer	**18**	**0**	**-**	**-**
(1 = yes, 2 = no)	*12*	*11*	-	-
15. Attribute speech improvements to	**16**	**0**	**0**	**-**
(1 = CSP, 2 = other therapy, 3 = factors other than therapy)	*18*	*0*	*3*	-
16. In order to be fluent I must pay attention to my speech	**13**	**4**	**7**	**0**
(1. almost always, 2. most of the time, 3. some of the time, 4. seldom)	*9*	*10*	*4*	*0*
17.My fluency skills 'work'	**7**	**10**	**5**	**1**
(1. almost always, 2. most of the time, 3. some of the time, 4. seldom)	*1*	*14*	*6*	*2*
18. As a result of CSP therapy confidence in my ability to speak is	**8**	**10**	**5**	**1**
(1. much improved, 2. moderately improved, 3. slightly improved, 4. not improved)	8	10	4	1
19. As a result of CSP therapy general confidence is	**10**	**7**	**5**	**2**
(1. much improved, 2. moderately improved, 3. slightly improved, 4. not improved)	*8*	*8*	*7*	*0*

3.3 The Doetinchem Method (DM)

3.3.1 Speech behaviour

3.3.1.1 Stutter severity

Number of syllables per minute (SPM)

In Table 3.6 the mean SPM and mean differences (averaged across the four speech tasks) relative to pretreatment are given for the DM group. There were no significant SPM differences between the pre- and post-treatment assessment, nor for follow-up 1 and follow-up 2.

Table 3.6. Mean SPM and %SS and differences between pre- and post-treatment, between pretreatment and Follow-up 1 and between pretreatment and Follow-up 2, averaged over the four speech tasks (interview, monologue, reading, telephone conversation) for the DM group.

Measurement session	*Mean and differences relative to pretreatment *p<0.05; **p<0.01)*			
	SPM		*%SS*	
	Mean	*Differences relative to pretreatment*	*Mean*	*Differences relative to pretreatment*
Pretreatment	138.2		8.6	
Post-treatment	162.5	-24.4	5.7	2.9
Follow-up 1	153.3	-15.1	6.2	2.4
Follow-up 2	150.0	-11.8	6.1	2.5*

In Figure 3.5 the mean SPMs for the speech tasks are shown. The differences between pre- and post-treatment measures were significant for the monologue ($p<0.05$). This effect was maintained in the F1 measure ($p<0.05$). The other differences did not reach significance. It should be noted, however, that our DM group consisted of only 14 participants (one client had decided not to participate in the F1 and F2 sessions). With fewer subjects, the statistical power of the DM analysis was reduced compared to the other therapy programmes.

Percentage of stuttered syllables (%SS)

Table 3.6 lists the overall mean %SS and mean differences between the four measurements. As mentioned above, the DM group had a smaller sample size (n=14) compared to the CSP group, resulting in a lower statistical power, which hampered the detection of possible treatment effects. Although the difference between the pre- and post-treatment measures was not significant, the mean %SS had decreased by 34% and this was maintained in F1 (28.3%) and F2 (29.9%). Thus, the gain was only significant in F2. Figure 3.6 shows the mean (and SE) %SS of the DM group per speech task for all measurements. The difference between pre- and post-treatment values was significant for the monologue task ($p<0.05$). At the one-year follow-up there was no longer any treatment effect. Two years after treatment there was a significant difference for the interview ($p<0.01$) as compared to the pretreatment measurement.

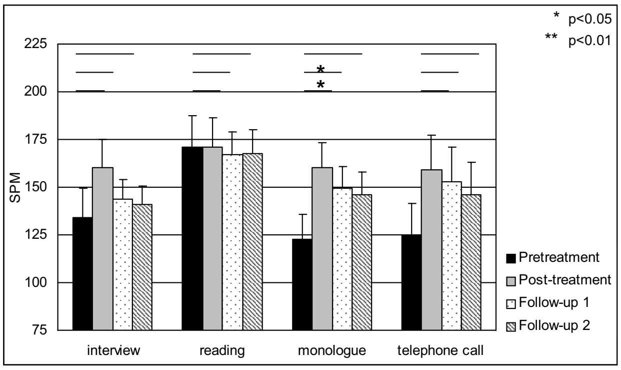


Figure 3.5. Mean number (and SE) of syllables per minute (SPM) for the DM group in the interview, reading task, monologue and telephone call for each measurement (pre, post, F1 and F2).

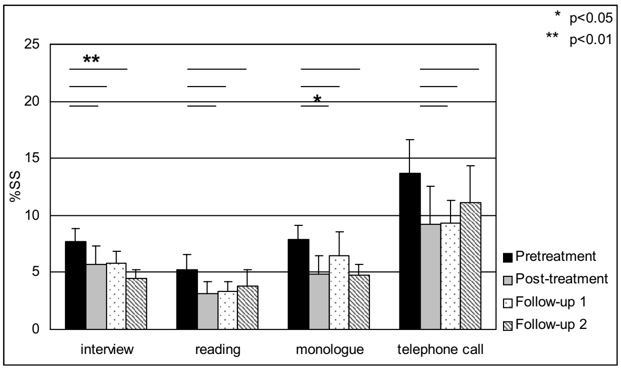


Figure 3.6. Mean percentage (and SE) of stuttered syllables (%SS) for the DM group in the interview, reading task, monologue and telephone call for each measurement (pre, post, F1 and F2).

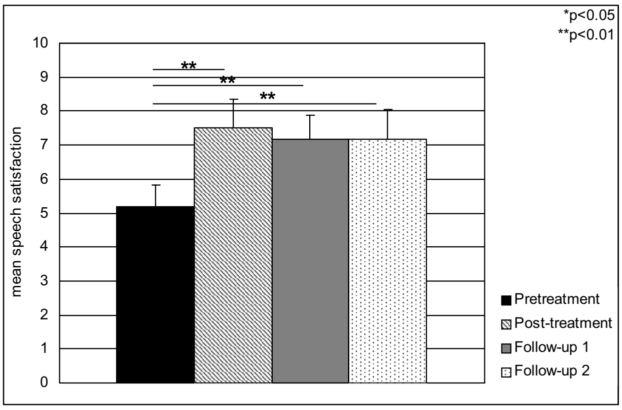


Figure 3.7. Mean scores (and SDs) on the speech satisfaction scale for each measurement of the DM group.

Self-evaluation

Figure 3.7 shows the mean scores on the satisfaction rating scale for the DM group. At pretreatment assessment, the mean rating of the participants was 5.18 (SD=0.66). The rating had significantly increased (p<0.01) to 7.50 (SD=0.83) immediately following therapy. In the long term there was some regression (7.18, SD=0.70 for F1 and 7.18, SD=0.87 for F2) but the improvements relative to the pretreatment means remained significant (p<0.01).

3.3.1.2 Speech quality

The Figures 3.8a, b and c reflect the mean scores of the DM group for each of the three factors. On the Voice Dynamics factor (Figure 3.8a), the score had decreased at post-treatment assessment and this effect was maintained in F1 and F2. Recall that a lower score on this factor signifies more dynamic speech. There was a significant difference between pre- and post-treatment means ($F(1,12)=12.079$; $p=0.005$), as well as between the pretreatment and F1 ($F(1,12)=11.806$; $p=0.005$) and pretreatment and F2 means ($F(1,12)=22.064$; $p=0.001$). Thus, the voice dynamics had improved as a result of therapy. Figure 3.8b shows the mean scores on the factor Articulation Quality. The higher the score on this factor the better the judgments were. The post-treatment value was lower than that of the pretreament assessment, but in F1 and F2 the mean scores increased again although the gains were not significant.

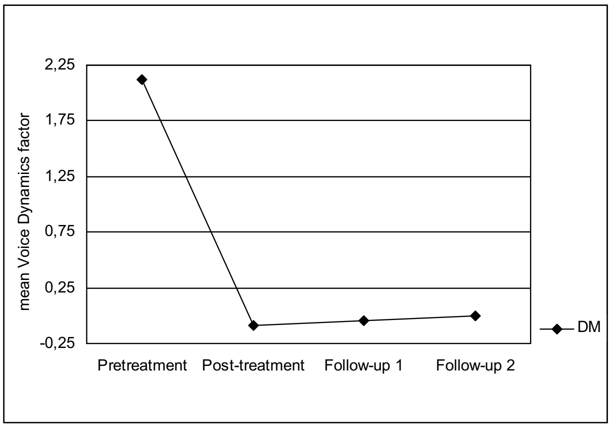


Figure 3.8a. Mean scores of the DM group on the Voice Dynamics factor for each assessment.

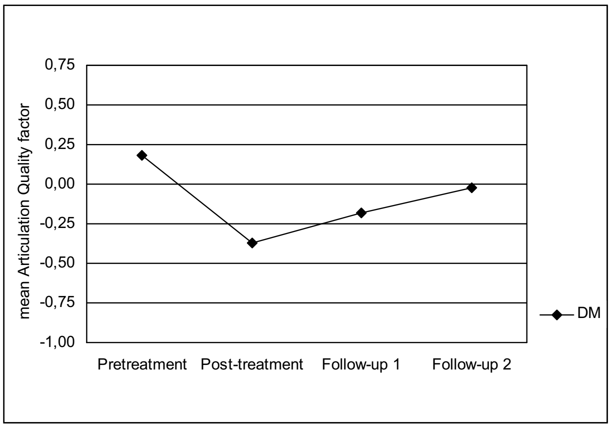


Figure 3.8b. Mean scores of the DM group on the Articulation Quality factor for each assessment.

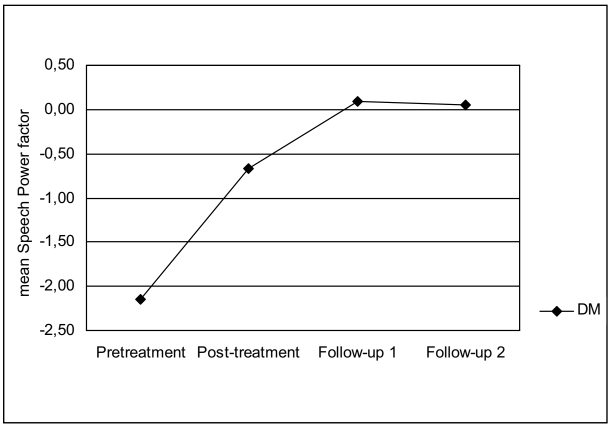


Figure 3.8c. Mean scores of the DM group on the Speech Power factor for each assessment (pre, post, F1 and F2).

Figure 3.8c depicts the mean scores for the Speech Power factor. The higher the score on this factor, the more powerful the speech. The mean score had increased significantly at post-treatment measurement ($F(1,12)=8.095$; $p=0.015$) and had improved even more in F1 ($F(1,12)=22.762$; $p=0.000$) and F2 ($F(1,12)=19.404$; $p=0.001$). Thus, as a result of therapy speech power was significantly enhanced, both in the short and in the long term.

3.3.1.3 Speech-motor control

Immediately after therapy there were no significant gains in the number of correctly produced syllables for any of the sequences. In F1 there was an overall slight increase and even in F2 the increment was only significant for the sequences /pə/ ($p<0.01$) and /pətəkə/ ($p<0.05$). The two-syllable sequence /təkə/ showed no effect of treatment. Table 3.7 lists the mean scores and mean differences relative to the pretreatment measures.

Table 3.7. Mean scores (number of correctly produced syllables) and differences relative to pretreatment for the DM group per sequence in the DDK task.

	*Mean scores and mean differences relative to pretreatment for the DDK (*p<0.05, **p<0.01)*					
	pə (SE)		*təkə (SE)*		*pətəkə (SE)*	
DM (n=11)	*Mean*	*Mean differences*	*Mean*	*Mean differences*	*Mean*	*Mean differences*
Pretreatment	20.41		20.05		14.92	
Post-treatment	19.18	1.23	19.09	0.96	16.08	-1.16
Follow-up 1	21.95	-1.54	21.00	-0.95	16.79	-1.87
Follow-up 2	23.23	-2.82**	20.50	-0.45	16.46	-1.54*

3.3.2 Self-report questionnaires

The responses to the self-report questionnaires are presented in Table 3.8. The mean differences between pre- and post-treatment scores are listed, as well as those for the first and second follow-up measures, both relative to the pretreatment measures. Significant differences are indicated with *(p<0.05) and **(p<0.01).

(1) Perceptions of Stuttering Inventory (PSI)

The mean PSI pretreatment scores for the DM group were: Struggle (S)=11.00 (SE=0.89); Avoidance (A)=7.30 (SE=1.50); and Expectancy (E)=9.70 (SE=1.28). Post-treatment struggle behaviour and avoidance had decreased (p<0.01) but this effect was not maintained in F1 and F2. Relative to the pretreatment measures, the expectancy scores had significantly decreased in F1 and F2 (p<0.05).

(2) Brutten's Speech Situation Checklist (SSC)

On the Emotional Reaction (ER) dimension the DM group had a mean pretreatment score of 146.0 (SD=29.4), which score had decreased significantly at post-treatment (p<0.05; 97.7, SD=50.7) as well as at F1 and F2 (90.3, SD=44.4; 106.0, SD=26.4, respectively). Although there was some regression relative to the pretreatment scores, the effect remained significant (p<0.05 and p<0.01, respectively). On the Distorted Speech (DS) dimension the DM participants showed a pretreatment mean of 136.3 (SD=23.9) and a significant decreased post-treatment mean (p<0.05; 87.9, SD=41.1). In F1 this effect had been maintained (p<0.05) but in F2 there was no longer any gain compared to the pretreatment mean.

(3) S24 and S44

At pretreatment evaluation, the DM group had a mean score of 15.10 (SE=1.80) on the S24 and 26.20 (SE=2.45) on the S44. Their post-treatment scores showed a significant decrease (p<0.01) to 8.80 (SE=1.09) for the S24 and to 15.30 (SE=1.57) for the S44. This effect was maintained (p<0.01) in F1 for both the S24 (9.80, SE=1.40) and the S44 (16.50, SE=1.97) but in F2 this applied to the S44 only (p<0.05).

(4) Lanyon's Stuttering severity scale

At the pretreatment evaluation the DM group showed a mean score of 41.55 (SE=3.12). At post-treatment assessment the score had decreased significantly (p<0.01) to 23.45 (SE=2.54). In F1 the mean score was 27.73 (SE=3.60) and in F2 this was 26.64 (SE=4.10), implying that the effect relative to pretreatment was maintained in both F1 and F2 (p<0.01) and, thus, that after therapy the participants' stutter severity had decreased in the given situations.

(5) Inventory of Interpersonal Situations (IIS)

Post-treatment tension scores revealed a significant decrease on three of the five subscales: 'Criticising' (p<0.05), 'Getting attention' (p<0.05), and 'Initiating a conversation' (p<0.01). The subscales 'Appreciation' and 'Appreciating yourself' had not been reduced significantly. These results were maintained in F1 and F2 for the subscales 'Criticising' (F1: p<0.01; F2: p<0.05), 'Getting attention' (F1 en F2: p<0.01) and for the subscale 'Initiating a conversation' (F1 en F2: p<0.01). On the Frequency scale there was only an increase in frequency for the subscale 'Initiating a conversation' ($p<0.05$), although in F1 and F2 this effect had disappeared (see Table 3.8).

(6) Hermans' Achievement Motivation Test (AMT)

The DM group had a mean pretreatment achievement motivation of 18.80 (SE=1.71), a mean negative failure anxiety of 16.00 (SE=0.47) and a mean positive failure anxiety of 8.20 (SE=1.05). There was no effect of therapy on the AMT scale. The negative anxiety failure had decreased at post-treatment assessment ($p<0.05$) whereas the positive anxiety failure had increased at F2 ($p<0.01$; also see Table 3.8).

Table 3.8. Mean differences (and standard errors in parentheses) between pre- and post-treatment, pretreatment and Follow-up 1 and between pretreatment and Follow-up 2 on the self-report questionnaires for the DM group.

Questionnaires	*Subscales*	*Mean Differences DM (*p<0.05, **p<0.01)*		
		Pre-Post (SE)	*Pre-F1(SE)*	*Pre-F2(SE)*
Perceptions of Stuttering Inventory (PSI)	*Struggle*	4.40 (1.08)**	2.30 (1.37)	3.60 (1.61)
	Avoidance	4.40 (1.19)**	2.20 (0.99)	2.90 (1.55)
	Expectancy	2.50 (1.23)	3.00 (0.92)*	3.40 (1.42)*
Brutten's SSC	*Emotional Reaction*	0.95 (0.36)*	1.09 (0.34)*	0.78 (0.17)**
	Distorted Speech	0.95 (0.27)*	0.79 (0.30)*	0.50 (0.21)
Erickson/ Andrews & Cutler	*S44*	10.90 (3.01)**	9.70 (2.57)**	9,60 (3,43)*
	S24	6.30 (1.80)**	5.30 (1.34)**	4,90 (2,46)
Lanyon's Stuttering Severity scale		18.09 (2.77)**	13.82 (2.40)**	14.91 (4.54)**
Inventory of Interpersonal Situation (IIS)	*Criticising*	4.44 (1.60)*	3.67 (0.99)**	4.22 (1.49)*
	Getting attention	4.56 (1.42)*	4.11 (0.96)**	6.11 (1.27)**
	Appreciation	1.00 (0.60)	0.22 (0.66)	1.11 (0.59)
	Initiating a conversation	5.33 (1.22)**	4.22 (0.89)**	5.67 (1.31)**
	Appreciating yourself	1.44 (1.12)	0.67 (0.78)	1.67 (0.90)
Tension	*Total*	21.56 (6.60)*	15.56 (3.88)**	24.00 (5.55)**
Inventory of Interpersonal Situation (IIS)	*Criticising*	-1.30 (1.49)	-0.70 (1.27)	-0.40 (1.74)
	Getting attention	-1.50 (1.49)	-2.10 (0.70)	-2.20 (1.64)
	Appreciation	-1.30 (0.67)	-0.40 (0.70)	-1.10 (0.75)
	Initiating a conversation	-2.80 (1.13)*	-1.30 (1.16)	-2.30 (1.19)
Frequency	*Appreciating yourself*	-2.60 (1.19)	-0.90 (1.24)	-1.40 (1.71)
	Total	-13.10 (5.74)*	-8.80 (5.85)	-11.20 (7.19)
Hermans' AMT	*Achievement Motivation (A)*	1.60 (1.69)	-0.40 (2.05)	0.50 (1.66)
	Neg. Failure Anxiety (F-)	3.20 (1.25)*	2.60 (1.16)	2.60 (1.32)
	Pos. Failure Anxiety (F+)	0.30 (0.83)	-0.70 (0.93)	-2.80 (0.76)**

Table 3.9. Number of responses per scale (1, 2, 3 and 4) for each item of the SPQ at F1 and F2 for the DM sample.

DM Responses to the SPQ *(F1: n=12; F2: n=11)*	*Number of responses* **F1** *and F2*			
	1	***2***	***3***	***4***
1. Satisfaction with speech before therapy	**0**	**2**	**8**	**2**
(1. very satisfied, 2. generally satisfied, 3. generally dissatisfied, 4. very dissatisfied)	*0*	*0*	*8*	*3*
2. Satisfaction with speech immediately after therapy	**4**	**7**	**1**	**0**
(1. very satisfied, 2. generally satisfied, 3. generally dissatisfied, 4. very dissatisfied)	*3*	*8*	*0*	*0*
3. Current rating of speech satisfaction	**1**	**7**	**3**	**1**
(1. very satisfied, 2. generally satisfied, 3. generally dissatisfied, 4. very dissatisfied)	*1*	*7*	*3*	*0*
4. Current rating of speech fluency	**0**	**9**	**3**	**0**
(1. very good, 2. generally good, 3. generally dissatisfied, 4. very dissatisfied)	*0*	*10*	*1*	*0*
5. Now have necessary skills to control speech	**2**	**5**	**5**	**0**
(1. almost always, 2. most of the time, 3. some of the time, 4. seldom)	*2*	*4*	*2*	*2*
6. Now have necessary skills to sound fluent	**2**	**5**	**4**	**1**
(1. almost always, 2. most of the time, 3. some of the time, 4. seldom)	*3*	*5*	*2*	*1*
7. Now have necessary skills to sound normal	**2**	**9**	**1**	**0**
(1. almost always, 2. most of the time, 3. some of the time, 4. seldom)	*3*	*5*	*3*	*0*
8. I use my speech controls	**2**	**2**	**7**	**1**
(1. almost always, 2. most of the time, 3. some of the time, 4. seldom)	*1*	*3*	*5*	*2*
9. Now able to speak normally without thinking about controls	**1**	**6**	**5**	**0**
(1. almost always, 2. most of the time, 3. some of the time, 4. seldom)	*0*	*7*	*2*	*2*
10. Now feel like a normal speaker	**2**	**6**	**4**	**0**
(1. almost always, 2. most of the time, 3. some of the time, 4. seldom)	0	8	2	1
11. As a result of therapy my speech fluency is	**3**	**4**	**3**	**1**
(1. much improved, 2. moderately improved, 3. slightly improved, 4. not improved)	*5*	*2*	*4*	*0*
12. I found the DM therapy programme to be	**8**	**2**	**2**	**0**
(1. very helpful, 2. moderately helpful, 3. slightly helpful, 4. not helpful)	*9*	*0*	*2*	*0*
13. Prefer stuttering over controlled speech	**0**	**0**	**3**	**9**
(1. all of the time, 2. most of the time, 3. some of the time, 4. seldom)	*0*	*4*	*1*	*6*
14. Currently consider myself a stutterer	**8**	**4**	**-**	**-**
(1 = yes, 2 = no)	*8*	*3*	-	-
15. Attribute speech improvements to*	**9**	**1**	**0**	**-**
(1 = DM, 2 = other therapy, 3 = factors other than therapy)	*8*	*1*	*1*	-
16. In order to be fluent I must pay attention to my speech	**5**	**3**	**3**	**1**
(1. almost always, 2. most of the time, 3. some of the time, 4. seldom)	*4*	*1*	*5*	*1*
17.My fluency skills 'work'	**2**	**8**	**2**	**0**
(1. almost always, 2. most of the time, 3. some of the time, 4. seldom)	2	6	2	1
18. As a result of DM therapy confidence in my ability to speak is	**6**	**5**	**1**	**0**
(1. much improved, 2. moderately improved, 3. slightly improved, 4. not improved)	*7*	*2*	*2*	*0*
19. As a result of DM therapy general confidence is	**8**	**3**	**1**	**0**
(1. much improved, 2. moderately improved, 3. slightly improved, 4. not improved)	*7*	*2*	*1*	*1*

Items 11 and 12 in F1 and items 5 and 15 in F2 were not completed by some of the participants.

The Speech Performance questionnaire (SPQ)

Table 3.9 lists for each item on the SPQ how often the particular score occurred in both F1 and F2. Seventy-five percent of the 12 clients in F1 and 91% of the 11 clients in F2 rated their current speech fluency as 'generally good' (item 4). Eighty-three percent of the 12 clients in F1 and 82% of the 11 clients in F2 found the DM 'very helpful' or 'moderately helpful' (item 12) and 90% of the 12 clients in F1 and 80% of the 11 clients in F2 attributed the speech improvements to the DM programme (item 15). Ninety-two percent of the 12 clients in F1 and 82% of the 11 clients in F2 indicated their confidence in speech as 'much improved' or 'moderately improved' (item 18).

3.4 The VSN individual programme

3.4.1 Speech behaviour

3.4.1.1 Stutter severity

Number of syllables per minute (SPM)

The mean SPM values had increased significantly ($p<0.05$) at post-treatment evaluation (21.9%) but due to regression the effect had disappeared at F1. At F2 the SPM scores had increased again, resulting in a significant ($p<0.05$) improvement of 20% relative to pretreatment (see Table 3.10).

In Figure 3.9 the mean number of syllables per minute are given for each speech task. Although there was an increased post-treatment number of syllables for all speech tasks, the effect was only significant for the monologue ($p<0.05$). There was a slight regression in F1.

Table 3.10. Mean SPM and %SS and differences between pre- and post-treatment, between pretreatment and Follow-up 1 and between pretreatment and Follow-up 2, averaged over the four speech tasks (interview, monologue, reading, telephone conversation) for the VSN group.

Measurement session	*Mean and differences relative to pretreatment *p<0.05; **p<0.01)*			
	SPM		*%SS*	
	Mean	*Differences relative to pretreatment*	*Mean*	*Differences relative to pretreatment*
Pretreatment	135.0		10.2	
Post-treatment	164.6	-29.6 *	6.3	3.9
Follow-up 1	158.6	-23.6	6.1	4.1
Follow-up 2	162.0	-27.0 *	6.1	4.1*

Percentage of stuttered syllables (%SS)

In Table 3.10 the overall means and mean differences between the four assessments are given. At post-treatment and at F1 the mean %SS had decreased compared to the pretreatment scores (38.2% and 39.9%, respectively) but this difference was not significant. Compared with the pretreatment measurement, the %SS at F2 again revealed a significant ($p<0.05$) reduction to 39.5%. As shown in Figure 3.10, there was a clear difference between pre- and post-treatment values. However, partly due to the small sample size (resulting from the large number of drop-outs) and large SD, these differences were not significant.

Self-evaluation

Figure 3.11 depicts the mean scores on the speech satisfaction rating scale for the VSN group. Prior to therapy, the mean rating of the participants was 4.88 (SD=0.81). Post-treatment, the means had increased significantly ($p<0.01$) to 6.84 (SD=1.00). In the longer term there was a slight regression (6.50, SD=0.87 for F1 and 5.92, SD=0.98 for F2) but, relative to the pretreatment scores, the gains remained significant ($p<0.01$).

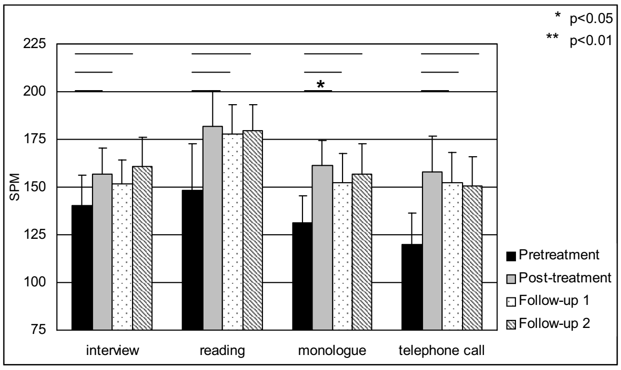


Figure 3.9. Mean number (and SE) of syllables per minute (SPM) for the VSN group in the interview, reading task, monologue and telephone call for each measurement (pre, post, F1 and F2).

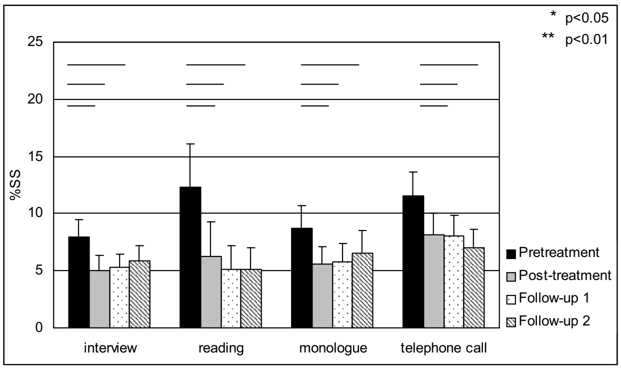


Figure 3.10. Mean percentage (and SE) of stuttered syllables (%SS) for the VSN group in the interview, reading task, monologue and telephone call for each measurement.

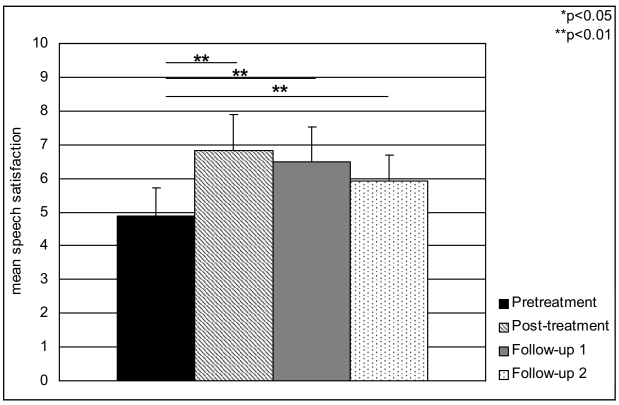


Figure 3.11. Mean scores (and SDs) on the speech satisfaction scale for the VSN group for each measurement.

3.4.1.2 Speech quality

The Figures 3.12a, b, and c depict the mean scores of the VSN group for each of the three factors. On the Voice Dynamics factor (Figure 3.12a), the score had decreased at post-treatment assessment (a lower score on this factor signifies more dynamic speech). The main effect *time* (pre, post, F1 and F2) approached significance ($F(3,36)=2.856$; $p=0.051$). The difference between pretreatment and F2 was significant ($F(1,12)=10.619$; $p=0.007$). Figure 3.12b shows the mean score on the factor Articulation Quality. The higher the score on this factor the better the judgements were. The main effect *time* (pre, post, F1 and F2) was significant ($F(3,36)=2.952$; $p=0.045$). The mean for post-treatment Articulation Quality had increased, but in F1 the mean score had dropped below the pretreatment level. F2 revealed a moderate increase but the mean score still remained below the pretreatment level. Thus, for Articulation Quality none of the gains were significant. Figure 3.12c reflects the mean score on the Speech Power factor. The higher the score on this factor, the more powerful the speech. The main effect *time* (pre, post, F1 and F2) was significant ($F(3,36)=8.867$; $p<0.0001$). The mean Speech Power had decreased at post-treatment but in F1 had significantly increased relative to the pretreatment score ($F(1,12)=5.626$; $p=0.035$), which also applied to F2 ($F(1,12)=5.726$; $p=0.034$). Thus, the Speech Power factor showed a significant improvement two years post therapy.

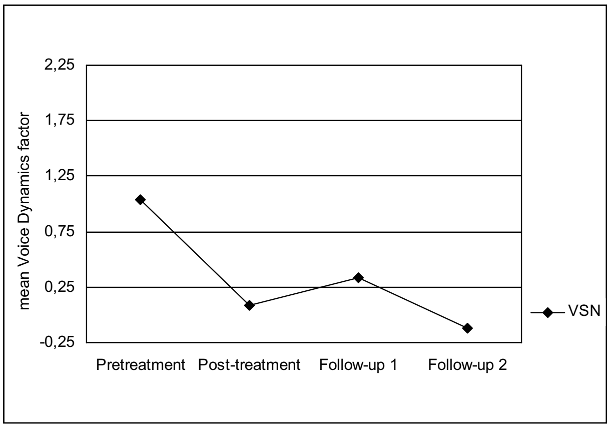


Figure 3.12a. Mean scores of the VSN group on the Voice Dynamics factor for each measurement.

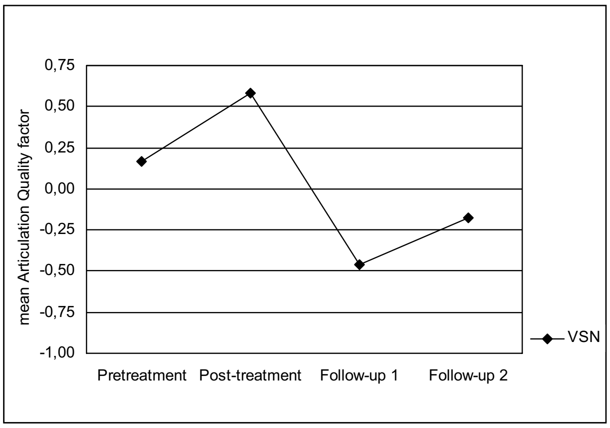


Figure 3.12b. Mean scores of the VSN group on the Articulation Quality factor for each measurement.

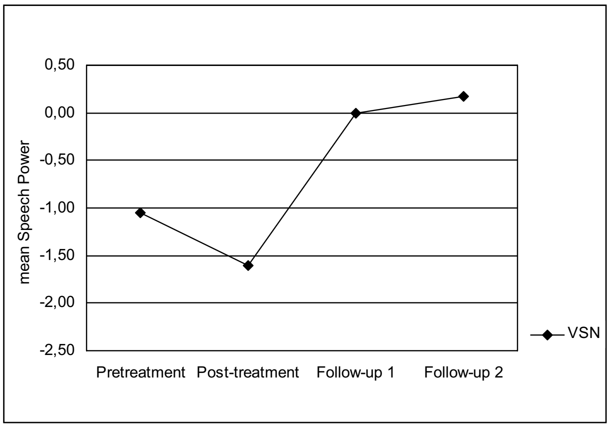


Figure 3.12c. Mean scores of the VSN group on the Speech Power factor for each measurement.

Table 3.11. Mean scores (number of correctly produced syllables) and mean differences relative to pretreatment for the VSN group per sequence in the DDK task.

	*Mean scores and mean differences relative to pretreatment for the DDK (*p<0.05, **p<0.01)*					
	/pə/ (SE)		*/təkə/ (SE)*		*/pətəkə/ (SE)*	
VSN (n=12)	*Mean*	*Mean differences*	*Mean*	*Mean differences*	*Mean*	*Mean differences*
Pretreatment	21.88		19.25		13.58	
Post-treatment	23.21	-1.33	20.92	-1.67	15.17	-1.59
Follow-up 1	25.00	-3.12	19.88	-0.63	15.00	-1.42
Follow-up 2	24.38	-2.50	19.42	-0.17	15.71	-2.13*

3.4.1.3 Speech-motor control

On the one- and two-syllable sequences there were no significant treatment effects. The only significant difference appeared on the three-syllable sequence /pətəkə/ at F2 ($p<0.05$). See also Table 3.11 for the mean scores and the mean differences.

3.4.2 Self-report questionnaires

The responses to the self-report questionnaires for the VSN sample are summarised in Table 3.12. The mean differences between pre- and post-treatment measures are provided as well as those between pretreatment and follow-up 1 and pretreatment and follow-up 2. Significant differences are indicated with * ($p<0.05$) and ** ($p<0.01$).

(1) Perceptions of Stuttering Inventory (PSI)

The mean pretreatment VSN scores on the PSI were: S=9.60 (SE=1.06), A=7.18 (SE=1.42) and E=8.82 (SE=1.45). The struggle behaviour (S) had decreased significantly in both F1 ($p<0.05$) and F2 ($p<0.01$). Post-treatment avoidance (A) had decreased significantly ($p<0.05$) and this effect was maintained in both F1 ($p<0.05$) and F2 ($p<0.01$). Stutter expectancy (E), however, only decreased in F2 ($p<0.05$).

(2) Brutten's Speech Situation Checklist (SSC)

At pretreatment measurement the VSN group had a mean score of 120.2 (SD=26.0) on the ER-dimension, which score had decreased at post-treatment assessment to 100.5 (SD=53.3). In the first follow-up the mean had further diminished to 91.0 (SD=50.5) but rose again in F2 to 110.8 (SD=10.8). Despite the decreased post-treatment mean score, the differences for emotional reaction were not significant relative to pretreatment. On the DS dimension the VSN group had a mean pretreatment score of 129.0 (SD=19.8) and a reduced post-treatment mean of 102.7 (SD=52.4), which decreased further in F1 to 96.8 (SD=51.4). However, it showed an increase to 113.2 (SD=23.4) in F2. Again, none of the differences were significant. Thus, neither the ER nor the DS dimension had been reduced as a result of therapy.

(3) S24 and S44

The pretreatment mean score was 14.70 (SE=1.12) for the S24 and 24.30 (SE=2.29) for the S44. Post-treatment S24 scores had decreased to 11.60 (SE=1.60) and the S44 mean had been reduced to 20.00 (SE=2.61). In F1 both the mean scores on the S24 and the S44 had increased, to 12.00 (SE=1.69) and 19.50 (SE=2.83), respectively. At F2 the mean score for the S24 had dropped again to 11.40 (SE=1.81) and for the S44 this was 19.10 (SE=2.90). Although the mean scores had decreased at the post-treatment evaluation, relative to the pretreatment means there was no significant difference for either the S24 or the S44.

(4) Lanyon's Stuttering Severity Scale

The mean VSN pretreatment score on the SS-scale was 36.10 (SE=3.26). There was a significantly lower ($p<0.01$) post-treatment mean of 28.30 (SE=4.50). At F1 the score had increased slightly (29.90; SE=4.91) but, relative to pretreatment, the effect was maintained ($p<0.05$). In F2 the score further decreased to 26.90 (SE=4.44), ($p<0.01$). Thus, both the short- and long-term stutter severity had decreased.

(5) Inventory of Interpersonal Situations (IIS)

At the post-treatment assessment none of the five subscales of the IIS revealed a significant decrease for tension. However, at F1 the tension for the subscale 'Initiating a conversation' had decreased significantly ($p<0.05$) and at F2 this applied to the two subscales 'Getting attention' ($p<0.05$) and 'Initiating a conversation' ($p<0.05$). On the frequency scale there was a significant difference between pre- and post-treatment scores on the subscale 'Appreciating yourself' ($p<0.01$). In F1 the only significant difference relative to pretreatment was found for the subscale 'Initiating a conversation' ($p<0.05$) and in F2 there was again an effect on the subscale 'Appreciating yourself' ($p<0.05$). Thus, the participants' post-treatment tension was lower when they needed to initiate a conversation and when they received attention. Furthermore, our sample more frequently initiated a conversation and appreciated themselves more often after therapy.

(6) Hermans' Achievement Motivation Test (AMT)

At pretreatment assessment the VSN group showed a mean achievement motivation of 17.80 (SE=2.71), a mean negative failure anxiety of 13.70 (SE=1.83) and a mean positive failure anxiety of 9.50 (SE=1.39). There was no effect of therapy on the AMT (see also Table 3.12).

Table 3.12. Mean differences (and standard errors in parentheses) between pre- and post-treatment, pretreatment and Follow-up 1 and between pretreatment and Follow-up 2 for the VSN group.

Questionnaires	*Subscales*	*Mean Differences (*p<0.05, **p<0.01)*					
		Pre-Post (SE)		*Pre-F1 (SE)*		*Pre-F2 (SE)*	
Perceptions of Stuttering Inventory (PSI)	*Struggle*	0.45	(0.95)	3.45	(1.27)*	3.00	(0.89)**
	Avoidance	3.36	(1.11)*	3.18	(1.18)*	4.09	(0.99)**
	Expectancy	1.18	(1.35)	1.64	(1.36)	3.00	(1.01)*
Brutten's SSC	*Emotional Reaction*	0.39	(0.43)	0.57	(0.40)	0.38	(0.18)
	Distorted Speech	0.52	(0.49)	0.63	(0.47)	0.31	(0.18)
Erickson/ Andrews &Cutler	*S44*	4.30	(2.16)	4.80	(2.74)	5.20	(2.93)
	S24	3.10	(1.55)	2.70	(1.70)	3.30	(1.95)
Lanyon's Stuttering Severity Scale		7.80	(1.97)**	6.20	(2.50)*	9.20	(1.70)**
Inventory of Interpersonal Situation (IIS) Tension	*Criticising*	0.36	(1.25)	0.45	(1.80)	0.73	(1.01)
	Getting attention	1.09	(0.89)	0.91	(1.36)	2.09	(0.77)*
	Appreciation	-1.09	(0.99)	0.36	(0.58)	0.18	(0.46)
	Initiating a conversation	1.00	(0.79)	2.55	(0.85)*	2.27	(0.79)*
	Appreciating yourself	-0.36	(0.45)	0.27	(0.84)	-0.09	(0.76)
	Total	0.27	(4.04)	6.00	(5.88)	6.18	(3.53)
Inventory of Interpersonal Situation (IIS) Frequency	*Criticising*	-2.64	(1.40)	-2.64	(2.03)	-2.55	(1.38)
	Getting attention	-1.55	(1.02)	-1.27	(1.35)	-0.91	(0.91)
	Appreciation	-0.73	(0.90)	-0.82	(1.14)	-0.73	(0.71)
	Initiating a conversation	-1.27	(0.63)	-2.09	(0.93)*	-1.27	(0.85)
	Appreciating yourself	-2.91	(0.79)**	-2.55	(1.21)	-2.91	(1.00)*
	Total	-11.18	(5.45)	-12.00	(7.42)	-12.09	(5.62)
Hermans' AMT	*Achievement Motivation (A)*	2.10	(1.25)	1.40	(2.02)	2.10	(1.90)
	Neg. Failure Anxiety (F-)	1.60	(1.01)	2.20	(1.00)	1.20	(0.63)
	Pos. Failure Anxiety (F+)	-1.40	(1.39)	-1.60	(1.28)	-0.60	(1.51)

Table 3.13. Number of responses per scale (1, 2, 3 and 4) for each item of the SPQ for the VSN sample.

VSN Responses to the SPQ (F1: n=11; F2: n=12)	*Number of responses* **F1** *and F2*			
	1	***2***	***3***	***4***
1. Satisfaction with speech before therapy	**1**	**2**	**7**	**1**
(1. very satisfied, 2. generally satisfied, 3. generally dissatisfied, 4. very dissatisfied)	0	1	7	4
2. Satisfaction with speech immediately after therapy	**1**	**9**	**0**	**1**
(1. very satisfied, 2. generally satisfied, 3. generally dissatisfied, 4. very dissatisfied)	1	8	3	0
3. Current rating of speech satisfaction	**2**	**7**	**0**	**2**
(1. very satisfied, 2. generally satisfied, 3. generally dissatisfied, 4. very dissatisfied)	0	8	3	1
4. Current rating of speech fluency	**0**	**8**	**3**	**0**
(1. very good, 2. generally good, 3. generally dissatisfied, 4. very dissatisfied)	0	8	3	1
5. Now have necessary skills to control speech	**2**	**4**	**3**	**2**
(1. almost always, 2. most of the time, 3. some of the time, 4. seldom)	1	6	4	1
6. Now have necessary skills to sound fluent	**1**	**5**	**3**	**1**
(1. almost always, 2. most of the time, 3. some of the time, 4. seldom)	3	3	5	1
7. Now have necessary skills to sound normal	**1**	**6**	**4**	**0**
(1. almost always, 2. most of the time, 3. some of the time, 4. seldom)	1	6	4	1
8. I use my speech controls	**0**	**5**	**5**	**1**
(1. almost always, 2. most of the time, 3. some of the time, 4. seldom)	2	3	7	0
9. Now able to speak normally without thinking about controls	**2**	**3**	**3**	**3**
(1. almost always, 2. most of the time, 3. some of the time, 4. seldom)	2	4	5	1
10. Now feel like a normal speaker	**3**	**3**	**3**	**2**
(1. almost always, 2. most of the time, 3. some of the time, 4. seldom)	3	4	2	3
11. As a result of therapy my speech fluency is	**3**	**4**	**3**	**1**
(1. much improved, 2. moderately improved, 3. slightly improved, 4. not improved)	3	6	2	1
12. I found the VSN therapy programme to be	**4**	**4**	**3**	**0**
(1.very helpful, 2. moderately helpful, 3. slightly helpful, 4. not helpful)	4	6	2	0
13. Prefer stuttering over controlled speech	**0**	**0**	**4**	**7**
(1. all of the time, 2. most of the time, 3. some of the time, 4. seldom)	0	0	2	10
14. Currently consider myself a stutterer	**8**	**0**	**-**	**-**
(1 = yes, 2 = no)	7	5	-	-
15. Attribute speech improvements to	**7**	**0**	**3**	**-**
(1 = VSN, 2 = other therapy, 3 = factors other than therapy)	6	0	6	-
16. In order to be fluent I must pay attention to my speech	**4**	**3**	**3**	**1**
(1. almost always, 2. most of the time, 3. some of the time, 4. seldom)	3	6	3	0
17.My fluency skills 'work'	**2**	**5**	**3**	**1**
(1. almost always, 2. most of the time, 3. some of the time, 4. seldom)	1	4	6	1
18. As a result of VSN therapy confidence in my ability to speak is	**5**	**1**	**5**	**0**
(1. much improved, 2. moderately improved, 3. slightly improved, 4. not improved)	6	3	2	1
19. As a result of VSN therapy general confidence is	**5**	**2**	**3**	**1**
(1. much improved, 2. moderately improved, 3. slightly improved, 4. not improved)	3	5	3	1

The Speech Performance questionnaire (SPQ)
Table 3.13 shows for each item on the SPQ how often the particular score featured in both F1 and F2. Seventy-three percent of the 11 clients in F1 and 67% of the 12 clients in F2 rated their current speech fluency as 'generally good' (item 4). Again, 73% of the 11 clients in F1 and 83% of the 12 clients in F2 found the VSN 'very helpful' or 'moderately helpful' (item 12) and 70% of the 11 clients in F1 and 50% of the 12 clients in F2 attributed their speech improvements to the VSN treatment (item 15). Fifty-five percent of the 11 clients in F1 and 75% of the 12 clients in F2 indicated their confidence in speech as 'much improved' or 'moderately improved' (item 18). These scores seem to be lower than the scores of the other two treatments.

4 Discussion and conclusions

As early as in 1939 Johnson concluded that there is no such thing as ***the*** method for treating stuttering. Prins (1997) stated on this topic: "The challenge for clinicians is to select methods that unite their expertise with the specific needs and complaints of their clients". In the introduction we already underlined the complex nature of stuttering, which makes the selection of the optimal individual programme all the more challenging. The reason why the phenomenon of stuttering is so obscure may be explained by the fact that two, perhaps interrelating, levels of human functioning are assumed to be involved: the physical aspect of speech production itself and/or the speaker's cognitive and emotional disposition. How or why stuttering manifests itself is, of course, speaker-dependent. It is therefore ill-advised to administer the same treatment to all stutterers. Rather, the choice for a specific treatment programme should be based on each individual stuttering profile. This is an inherent feature of the VSN programme evaluated in the present study, since it is based on an individualised approach. Although in the two group therapies, *i.e.* the CSP and the DM, the whole group underwent the same treatment, adjustments were made to accommodate the individual needs of the clients.

A common problem in stuttering therapy is related to its long-term effects. Clients often show regression (or even complete relapse) one or two years after treatment termination. Several researchers have tried to formulate criteria or definitions for the assessment of relapse (*e.g.* Boberg, 1981; Craig, Franklin, & Andrews, 1984; Blood, 1995) but most of these criteria only relate to the proportion of stuttered syllables (%SS). Because neither the DM nor the VSN focus on fluency reduction in terms of %SS, these criteria will not be discussed here. We will restrict ourselves to the discussion of the short- and long-term effects for each of the variables.

Although it would be interesting to compare the three therapy programmes, this was not part of the research goal of the present study. Moreover, due to the unequal numbers of participants in the study samples, the statistical power varied for each group, which would have made a comparison unreliable. In addition, the ages, educational levels but also stutter severity of the participants proved to differ per group. As detailed in subsection 2.4, the participants in the DM group were younger than the participants of both the VSN and the CSP samples. Clearly, younger participants are in a different phase of their lives and may therefore have had different treatment needs than the older participants. This, then, may have affected the impact of the treatment. Also, the educational levels varied across samples. Overall, the CSP group had received more advanced education than the DM and VSN. It is likely that this difference may have affected the treatment results. However, since the reason for the disproportion was not further investigated any explanation as to the nature of such a potential effect would remain highly speculative.

Nevertheless, evaluation per treatment also improves our understanding of therapy effects. The three treatment programmes we examined differ both in their goals and in their approaches and we will therefore consider the results in this discussion in relation to the specific characteristics of each therapy content. Although all three treatment programmes have a broad description of goals, procedures, argumentations and reasoning, each programme can be characterised by one or two spearheads. The CSP can be typified as a therapy in which fluency-enhancing techniques are taught within the contexts of prolonged speech. The DM can be described as a therapy that reduces the negative factors that maintain stuttering and the VSN as an individualised therapy in which the content of the therapy is adjusted to the individual needs of the stutterer. Following a summary of the main treatment results, the outcomes will be discussed in the light of these characterisations.

The Comprehensive Stuttering Programme (CSP)

Speech behaviour: Post therapy, the participants' speech fluency had considerably improved on all speech tasks as was reflected by the higher speech rates in the interview, the monologue and telephone conversation and the superior self-evaluation ratings on the speech satisfaction scale. With respect to speech quality, there was no treatment effect on the factors voice dynamics and speech power, but the articulation quality had improved significantly. Furthermore, we found a higher number of correctly produced syllables in the sequence /pətəkə/ (maximum capacity). At *Follow-up 1* substantial improvement in fluency, speech rate, and self-evaluation and articulation quality were maintained in spite of regression relative to post-treatment measures. The maximum capacity of correctly produced syllables in /pətəkə/ remained higher compared to the pretreatment results. At *Follow-up 2* these gains were maintained.

Self-reports: Post therapy, the results revealed improvements of stuttering severity, speech anxiety and attitude. Most of the significant post-treatment gains were maintained in both *Follow-up 1* and *Follow-up 2* and 83% of the participants found the CSP very or moderately helpful.

Discussion

The most striking result of the CSP programme was the dramatic decrease of dysfluencies. Participants spoke almost completely fluently immediately after treatment. This effect is in line with the content and purpose of the programme. More than 70% of the time was spent on speech-motor control (see introduction) and this clearly resulted in a greatly enhanced fluency. It should be noted, however, that because the CSP group exhibited a high percentage of dysfluencies prior to treatment, this could account for part of the post-treatment gains. Nevertheless, the finding of the extraordinarily high percentage of clients with a positive effect should mostly be attributed to the intervention.

Although the CSP enhances fluency within the context of prolonged speech, the speech rate had also increased at the post-treatment assessment. It is possible that the increase was the result of a decrease in dysfluencies rather than of an essential increase in speech rate (it should be noted that repetitions within a stutter were not counted for the SPM). The fact that almost any novel manner of speaking (*e.g.* prolonging speech, changing speech rate, paying attention to breathing patterns) results in improved fluency (Bloodstein, 1995, p. 407) and that the end-products of many such treatments often are complaints of unnaturalness, lack of stability, a resistance to using the technique and relapse to the pre-therapy level (Dayalu & Kalinowski, 2002, p88), requires further discussion of the speech naturalness data and the maintenance of the gains achieved. It should be noted, that at the post-treatment evaluation the number of fluently spoken syllables in the DDK task had increased, showing an overall improved speech capacity in the stutterers. This effect was maintained in the long term. Furthermore, although most clients showed some regression relative to the highly enhanced post-treatment scores, compared to the pretreatment levels, treatment effects were maintained in the longer term. The most prominent relapse was found in the number of dysfluencies, and this was reflected in the speech satisfaction ratings. Nevertheless, despite the relapse, the participants still rated their speech significantly higher than at the pretreatment session, indicating satisfaction with the results achieved. Moreover, the articulation quality improved even further as was shown by the F1 and F2 evaluations. This implies that the naïve listeners, when judging the speech samples of the four different measurement sessions, had not noticed the relapse in, for example, the %SS. The self-report questionnaires showed a decrease in perceived stutter severity and speech anxiety and improved attitudes in both the intermediate and long-term follow-up.

It is remarkable that there were no drop-outs in the CSP group. This is possibly due to the group dynamics, which might have positively influenced the treatment outcome.

The Doetinchem Method (DM)

Speech behaviour: Post therapy we found improved speech fluency and higher speech rates on the monologue task and enhanced self-evaluation scores on the speech satisfaction scale. There was no treatment effect for the articulation quality factor, but both the voice dynamics and the speech power factor had improved significantly. There was a higher number of correctly produced syllables in the DDK sequence /pətəkə/. At *Follow-up 1,* in spite of regression relative to post-treatment measures, the gains in speech rate and self-evaluation scores were maintained. Due to the moderate regression, the statistically significant effect was not maintained with respect to the fluency measure (%SS). The enhancements in the voice dynamics and speech power were sustained. The maximum capacity of correctly produced syllables in the /pətəkə/ sequence remained higher compared to the pretreatment capacity. At *Follow-up 2*, again relative to pretreatment measures, the speech fluency in the interview had increased. The superior self-evaluation ratings and the improvements on the voice dynamics and speech power factor were maintained. There were no further improvements in the number of correctly produced syllables in the /pətəkə/ DDK sequence.

Self-reports: Post therapy, the scores on the questionnaires indicated improvements with respect to stuttering severity, speech anxiety and attitude. At *Follow-up 1*, relative to pretreatment, the statistically significant improvements were maintained on almost all questionnaires. There was a slight regression on PSI (S and A) and IIS scales. At

Follow-up 2, relative to pretreatment, statistically significant improvements were maintained on almost all questionnaires. There was a slight regression on the Brutten (DS), S24 item. More than 80% of the participants found the DM programme to be very or moderately helpful.

Discussion
The results for the DM can be summarised as having yielded an enhanced self-perception directly after the end of therapy. These outcomes were rather robust and were maintained in the long term. This is in line with the DM's focus on the reduction of negative factors that maintain stuttering. The highest gain was achieved on the speech satisfaction scale and on the self-report questionnaires. Prior to treatment, the DM group's mean speech satisfaction rating was slightly above the mean scores of the CSP and the VSN group. Although relative to pretreatment the CSP group had improved more immediately after the end of therapy, the long-term gain was highest in the DM group. This was due to the fact that the DM clients experienced hardly any regression (the score remained above 7) in contrast to the significant regression in the CSP group.
Compared to the CSP programme, speech fluency benefited less from the DM treatment, although it should be noted that the DM sample showed fewer pretreatment dysfluencies compared to the CSP group and that, consequently, there was less to be gained. The moderate decrease in post-treatment dysfluencies was not sustained in the long term. At F1, only the monologue showed an increased speech rate. On the other hand, the clients' speech had improved as far as the voice dynamics and speech power were concerned. Hence, according to the naïve listeners, the post-treatment voice dynamics (including pitch, speech melody, expression, rate and deepness) and the speech power (including high loadings with respect to loudness, power and accentuation) were superior compared to the pretreatment speech samples. This implies that the DM clients' speech sounded more dynamical and more powerful and, again, these effects were rather robust. Neither of the scales revealed a relapse in F1 and F2. That there was no effect for the articulation quality factor is in line with the fact that the DM mainly focuses on the emotions and cognitions that are related to the stuttering problem and that, consequently, less treatment time is allocated to advance speech fluency and articulation quality. Again, it should be noted that the DM sample was small and that one client had dropped out after the post-treatment assessment, resulting in a lower statistical power of the tests.

The VSN individual programme
Speech behaviour: Post therapy, the mean %SS had decreased but the difference relative to the pretreatment proportion was not significant. The results did show an increased speech rate, albeit only for the monologue, and an improved self-evaluation score on the speech satisfaction scale. There was no improvement on the voice dynamics factor, nor on the articulation quality or the speech power factor and no increase in correctly produced syllables in the DDK sequences. At *Follow-up 1* there were no further fluency and speech-rate improvements relative to pretreatment. In spite of a slight regression relative to post-treatment, statistically significant improvements in the self-evaluation ratings were maintained. Relative to pretreatment, the clients' speech power had increased. There were no improvements on the voice dynamics factor or the

articulation quality factor and no increase in the number of correctly produced syllables in the DDK sequences. At *Follow-up 2*, fluency was enhanced and speech rate had increased compared to the pretreatment levels. In spite of regression relative to the post-treatment measures, improvements in the self-evaluation scores were maintained. Both the voice dynamics and the speech power factor had improved significantly. The number of correctly produced syllables in the DDK sequences /pətəkə/ had also risen.

Self-reports: Post therapy, there was little improvement on the questionnaire ratings with respect to stutter severity, speech anxiety and attitude. At *Follow-up 1*, relative to post-treatment there was a gain on the S item of the PSI and also for initiating a conversation (subscale of the IIS). At *Follow-up 2* there were also improvements on the PSI and IIS, relative to pretreatment. More than 70% of the participants in F1 and more than 80% of the participants in F2 found the VSN programme to be very or moderately helpful.

Discussion

The most obvious result for the VSN group is the enhanced self-perception, in particular on the speech satisfaction scale. Although the mean post-treatment percentage of stuttered syllables had reduced, the difference was not significant. This was caused by the small sample size, the relatively large standard deviation but also, compared to the CSP sample, by a smaller reduction in post-treatment dysfluencies. In Figure 3.10, a decreased post-treatment %SS is shown and this level of dysfluencies was maintained in F1 and F2. Of course there were large individual differences in treatment content, but overall (see Table 2.2 in the method section), the therapy content of the VSN group focused more on the emotions and cognitions (67.1%) than on speech-motor control (32.9%). This may explain the lack of dysfluency reduction. Post-treatment speech satisfaction had improved compared to pre-therapy levels and, even though there was some regression in the follow-up evaluations and the decrease in the proportion of dysfluencies was limited, in the long term speech satisfaction levels remained higher than the levels that we found in the pretreatment assessment.
Immediately after treatment, the judgments for the factors voice dynamics, articulation and speech power had not risen. The positive effects did show up in the follow-up assessments: in F2 the voice dynamics had improved and in F1 and F2 the speech power proved to have increased.

Similar effects were found on the DDK task. There were no post-treatment or F1 effects but in F2 the articulation capacity had improved for the /pətəkə/ sequence.
Why these effects did not occur directly after the end of the treatment is unclear but the phenomenon may be associated with the prolonged treatment duration of individualised therapy. Group therapy is clearly defined with respect to both content and duration, whereas individualised therapy is not. Clients do not know when the treatment will end and there is always a gap of at least some days between two sessions. This, in combination with the lack of group cohesion, may account for the less spectacular gains for the VSN programme compared to those achieved with the CSP and DM interventions.

Conclusions

Each of the three therapy programmes evaluated in the present study proved to be effective but regression or relapse with respect to the percentage of stuttered syllables and also with respect to speech satisfaction appeared to be inevitable. Several researchers have defined criteria for relapse (see e.g. Boberg, 1981; Craig, Franklin, & Andrews, 1984; Blood, 1995) but, since most of these criteria only relate to the percentage of stuttered syllables (which is not the main goal in the DM and VSN therapy), it would not have been appropriate to apply these criteria in our context. Overall, even though each programme showed some relapse, treatment effects were maintained in the long term.

Our study did not confirm that speech therapy often results in unnatural speech (*e.g.* Dayalu and Kalinowski, 2002). Moreover, significant improvements on the articulation quality factor (which is related to and includes speech naturalness) were found for the CSP even though this intervention targets the modification of speech output in particular by using stuttering reducing techniques. It is argued that specifically these kinds of approaches often result in unnatural speech (see for a review Dayalu and Kalinowsky, 2002).

Our evaluations revealed that the outcomes for the individualised intervention were less spectacular than the effects yielded by the two group therapies. Nevertheless, the fact that the therapy content can be fine-tuned to each individual stutterer makes individual therapy particularly suitable for clients with severe emotional and cognitive problems. In individualised therapy the content as well as the client-therapist relation is not subject to change and, thus, clients that have (temporarily) stopped attending their sessions can easily resume their treatment whenever they wish. This can never be achieved in group therapy since group composition and dynamics vary per group.

In conclusion, we recommend an independent diagnostic examination prior to therapy, in which the client-specific physical, cognitive and emotional aspects of stuttering are assessed in detail. This is already the standard procedure for the VSN treatment programme. Such a diagnostic procedure allows the therapist to choose the most appropriate treatment for each individual stutterer, which may prevent or reduce the number of clients 'shopping around' for the best therapy. Since individual and group therapies offer a variety of potential remedies, it is advisable to provide several approaches within one treatment facility, as is the case with the VSN and DM programmes in the so-called Integraal Zorg Traject (IZT; integrated care programme), to thus afford the best prospects for speakers seeking treatment for their speech problems.

References

Ackermann, H., Hertrich, I., & Hehr, T. (1995). Oral Diadochokinesis in Neurological Dysarthrias. *Folia Phoniatrica et Logopaedica, 47,* 15-23.

Ackermann, H., Konczak, J., & Hertrich, I. (1997). The temporal control of repetitive articulatory movements in Parkinson's disease. *Brain and Language, 56,* 312-319.

Alfonso, P.J., & Van Lieshout, P.H.H.M. (1997). Stability and Flexibility in Speech Motor Control: Evidence from Perceptually Fluent Utterances of Persons who do and do not Stutter. In E.C. Healey & H.F.M. Peters (Eds.), (pp. 58-64). Nijmegen: Nijmegen University Press.

Andrews, G., & Cutler, J. (1974). Stuttering therapy: The relation between changes in symptom level and attitudes. Journal of Speech and Hearing Disorders. *Journal of Speech and Hearing Disorders, 39,* 312-319.

Ansel, B.M. (2004). Treatment efficacy research in stuttering. *Journal of Fluency Disorders, 18,* 121-123.

Blood, G.W. (1995). A behavioral-cognitive therapy program for adults who stutter: Computers and counselling. *Journal of Communication Disorders 28*, 165-180.

Bloodstein, O. (1969). *Handbook on Stuttering.* Chicago: National Easter Seal Society for Crippled Children and Adults.

Bloodstein, O. (1995). *A handbook on stuttering, Fifth edition.* San Diego: Singular Publishing Group, Inc.

Boberg, E., & Kully, D. (1994). Long-term results of an intensive treatment program for adults and adolescents who stutter. *Journal of Speech and Hearing Research, 37,* 1050-1059.

Brutten, G.J. (1975). Stuttering: Topography, assessment and behavior change strategies. In J. Eisenson (Ed.), *Stuttering: A second symposium.* New York: Harper & Row.

Brutten, G.J., & Janssen, P. (1981). A normative and factor analysis study of the responses of Dutch and American stutterers to the Speech Situations Checklist. In Washington DC: American Speech-Language-Hearing Association.

Conture, E.G., & Guitar, B. (1993). Evaluating efficacy of treatment of stuttering. *Journal of Fluency Disorders, 18,* 253-287.

Conture, E.G. (1996). Treatment efficacy: Stuttering. *Journal of Speech and Hearing Research, 39,* 518-526.

Cordes, A.K., & Ingham, R.J. (1998). *Treatment efficacy for stuttering.* San Diego: Singular Publishing Group, Inc.

Craig, A.R., Franklin, J., & Andrews, G. (1984). A scale to measure locus of control of behaviour. *British Journal of Medical Psychology 57*, 173-180.

Dayalu, V.N., & Kalinowski, J. (2002). Pseudofluency in adults who stutter: the illusory outcome of therapy. *Perceptual and Motor Skills, 94,* 96.

De Geus, E. (2000). To create learning communities of specialists in stuttering therapy: a review of the development and accomplishments of stuttering centres in Holland. IFA-presentation 2000.

De Geus, E., & Putker D. (2003). Action learning in de logopedie. Logopedie en Foniatrie, maart 2003.

De Nil, L.F. (1999). Stuttering: A neurophysiological perspective. In N. Bernstein Ratner & E.C. Healey (Eds.), *Stuttering research and practice: Bridging the gap* (pp. 85-102). Mahwah, New Jersey: Lawrence Erlbaum Associates.

De Nil, L.F., & Bosshardt, H.G. (2001). Studying stuttering from a neurological and cognitive information processing perspective. In H.G. Bosshardt, J.S. Yaruss, & H.F.M. Peters (Eds.), *Stuttering: research, Therapy and Self-Help. Proceedings of the Third World Congress on Fluency Disorders* (pp. 53-58). Nijmegen: Nijmegen University Press.

Diepstra, H.D., Huinck, W.J., Hulstijn, W., & Peters, H.F.M. (2001). The effect of word length on speech reaction time and word duration in the fluent word production of stuttering and nonstuttering subjects. In B. Maassen, W. Hulstijn, R.D. Kent, H.F.M. Peters, & P.H.H.M. Van Lieshout (Eds.), *Speech motor control in normal and disordered speech. 4th International Speech Motor Conference* (pp. 245-248). Nijmegen, the Netherlands: Van Tilt.

Erikson, C.W. (1969). Assessing communication attitudes among stutterers. *Journal of Speech and Hearing Research, 12,* 711-724.

Franken, M.C. (1997). *Evaluation of stuttering therapy. Development of tools for measuring speech quality.* Dissertation University of Nijmegen.

Franken, M.C., Boves, L., Peters, H.F.M., & Webster, R.L. (1995). Perceptual rating instrument for speech evaluation of stuttering treatment. *Journal of Speech and Hearing Research, 38,* 280-288.

Franken, M.C., Boves, L., & Peters, H.F.M. (1997). Evaluation of Dutch Precision Fluency-shaping Program. In E.C. Healey & H.F.M. Peters (Eds.), (pp. 303-308). San Francisco: Nijmegen University Press.

Frattali, C.M. (1998). Outcomes measurement: Definitions, dimensions, and perspectives. In C.M. Frattali (Ed.), *Measuring outcomes in speech-language pathology*. New York: Thieme.

Guitar, B. (1998). *Stuttering. An integrated approach to its nature and treatment.* (2nd ed.) Philadelphia: Lippincott Williams & Wilkins.

Hermans, H.J.M. (1968). *Handleiding bij de Prestatie Motivatie Test.* Amsterdam: Swets & Zeitlinger.

Hermans, H.J.M. (1970). A questionnaire measure of achievement motivation. *Journal of Applied Psychology, 54,* 353-363.

Huinck, W.J., Wouters, E.H.A., Hulstijn, W., & Peters, H.F.M. (2001). Differences between stuttering and non-stuttering people: Comparing diadochokinesis, sentences repetition and reaction time tasks. In B. Maassen, W. Hulstijn, R.D. Kent, H.F.M. Peters, & P.H. H.M. Van Lieshout (Eds.), *Speech motor control in normal and disordered speech* (pp. 311-315). Nijmegen, the Netherlands: Vantilt.

Janse, E. (2004). Word perception in fast speech: artificially time-compressed vs. naturally produced fast speech. *Speech Communication, 42,* 155-173.

Janssen, P. (1973). *Vergelijking van effecten van twee stotterbehandelingen.* Praeventiefonds onderzoek nr. 28-33.

Janssen, P. (1994). *Helpt de Doetinchemse Methode?* Stichting Stottertherapie Doetinchemse Methode.

Janssen, P. (1997). Etiologie van stotteren. Theorieën en modellen. Houten/Diemen, Bohn Stafleu van Loghum. Handboek Stem- Spraak- Taal Pathologie.

Johnson, W. (1939). The treatment of stuttering. *Journal of Speech Disorders, 4,* 170-173.

Kalinowski, A. G., Stuart, A., Sark, S., & Armson, J. (1996). Stuttering amelioration at various auditory feedback delays and speech rates. *European Journal of Disorders of Communication, 31,* 259-269.

Kraaimaat, F.W., Vanryckegem, M., & Van Dam-Baggen, R. (2002). Stuttering and social anxiety. *Journal of Fluency Disorders 27*, 319-331.

Kully, D. & Langevin, M. (1999). Intensive treatment for stuttering adolescents. In R.F.Curlee (Ed.), *Stuttering and related disorders of fluency* (2nd ed., pp. 139-159). New York: Thieme Medical Publishers.

Langevin, M. & Boberg, E. (1993). Comprehensive stuttering program. *Logopedie en Foniatrie, 10,* 262-268.

Langevin, M. & Boberg, E. (1993). Results of an intensive stuttering therapy program. *JSLPA, 17,* 158-166.

Lanyon, R.I. (1967). The measurement of stutter severity. *Journal of Speech and Hearing Research, 10,* 836-843.

Osgood, C.E., Succi, G.J., & Tannenbaum, P.H. (1957). *The measurement of meaning.* Urbana: University of Illinois Press.

Perkins, W.H. (1981). measurement and maintenance of fluency. In E. Boberg (Ed.), *Maintenance of fluency* (pp. 147-178). New York: Elsevier.

Peters, H.F.M. (1999). Stotteren: De spraakmotorische component. In E. Manders, D. Lembrechts, & P. Bastijns (Eds.), *Stotteren, recente inzichten* (pp. 93-117). Leuven.

Prins, D. (1997). Modifying stuttering- the stutterer's reactive behavior: perspectives on past, present, and future. In R.F. Curlee & G. M. Siegel (Eds.), *Nature and treatment of stuttering. New directions* (2nd ed., pp. 335-355). Boston: Allyn and Bacon.

Schoenaker, T. & Schoenaker, T. (1975). *Stottertherapie.* Züntersbach: Rudolf Dreikers Institut für soziale gleichwertigkeit.

Sacket, D.L., Rosenberg, W.M.C., Gray, J.A.M., Haynes, R.B., & Richardson, W.S. (1996). Evidence based medicine: what it is and what it isn't. *British medical journal, 312,* 71-72

Shrout, P.E. & Fleiss, J.L. (1979). Intra class correlations: Uses in assessing rater reliability. *Pschological Bulletin, 2,* 420-428.

Starkweather, C. W. (1993). Issues in the efficacy of treatment for fluency disorders. *Journal of Fluency Disorders, 18,* 151-168.

Thomas, C. & Howell, P. (2001). Assessing efficacy of stuttering treatments. *Journal of Fluency Disorders, 26,* 311-333.

Van Alphen, W. & Hoogerwerf, N. (1984*). Stotteren, helpt de Doetinchemse Methode? Een evaluatie-onderzoek naar resultaten van stottertherapie volgens de Doetinchemse Methode.* Stichting tot voorlichting over stotteren.

Van Dam-Baggen, C.M.J. & Kraaimaat, F.W. (1987). *Inventarisatielijst Omgaan met Anderen.* Lisse: Swets & Zeitlinger b.v.

Van Lieshout, P.H.H.M. (1995). *Motor planning and articulation in fluent speech of stutterers and nonstutterers.* Dissertation University of Nijmegen.

Van Lieshout, P.H.H.M., Hulstijn, W., & Peters, H.F.M. (1996a). Speech production in people who stutter: Testing the motor plan assembly hypothesis. *Journal of Speech and Hearing Research, 39,* 76-92.

Van Lieshout, P.H.H.M., Hulstijn, W., & Peters, H.F.M. (1996b). From planning to articulation in speech production: what differentiates a person who stutters from a person who does not stutter? *Journal of Speech and Hearing Research, 39,* 546-564.

Van Lieshout, P.H.H.M., Peters, H.F.M. & Bakker, A.J. (1997). En route to a speech motor test: A first halt. In W. Hulstijn, H.F.M. Peters, P.H.H.M. Van Lieshout (ed.) Speech production: *Motor control, brain research and fluency disorders*. Amsterdam: Elsevier Science.

Van Lieshout, P.H.H.M., Starkweather, C.W., Hulstijn, W., & Peters, H.F.M. (1995). Effects of linguistic correlates of stuttering on EMG activity in nonstuttering speakers. *Journal of Speech and Hearing Research, 38,* 360-372.

Van Riper, C. (1973). The treatment of stuttering. Englewood Cliffs: Prentice Hall.

Watson, B.C. & Alfonso, P.J. (1982). A comparison of LRT and VOT values between stutterers and nonstutterers. *Journal of Fluency Disorders, 7,* 219-241.

Watson, B.C. & Alfonso, P.J. (1983). Foreperiod and stuttering severity effects on acoustic laryngeal reaction time. *Journal of Fluency Disorders, 8,* 183-205.

Watson, B.C. & Alfonso, P.J. (1987). Physiological bases of acoustic LRT in nonstutterers, mild stutterers, and severe stutterers. *Journal of Speech and Hearing Research, 30,* 434-447.

Woolf, G. (1967). The assessment of stuttering as struggle, avoidance and expectancy. *British Journal of Disorders of Communication, 2,* 158-171.

Yairi, E. (1999). Epidemiologic factors and stuttering research. In N. Bernstein Ratner & E.C. Healey (Eds.), *Stuttering research and practice: Bridging the gap* (pp. 45-53). Mahwah, New Jersey: Lawrence Erlbaum Associates.

Yairi, E. & Ambrose, N.G. (1999). Early childhood stuttering I: Persistency and recovery rates. *Journal of Speech, Language, and Hearing Research, 42,* 1097-1112.

Yairi, E., Ambrose, N.G., & Cox, N.J. (1993). Genetics of stuttering: a critical review. *Journal of Speech, Language, and Hearing Research, 39,* 771-784.

Chapter Three

The relationship between pre-treatment clinical profile and treatment outcome in an integrated stuttering program

Huinck, W.J., Langevin, M., Kully, D., Graamans, K., Peters, H.F.M., & Hulstijn W. (2006). Journal of Fluency Disorders, 31, 43-63.

Abstract

A procedure for subtyping individuals who stutter and its relationship to treatment outcome is explored. Twenty-five adult participants of the Comprehensive Stuttering Program (CSP) were classified according to: (1) stuttering severity and (2) severity of negative emotions and cognitions associated with their speech problem. Speech characteristics (percentage of stuttered syllables, distorted speech score, and the number of correctly produced syllables on a diadochokinesis task) and emotional/ cognitive states (emotional reaction, speech satisfaction, and attitudes toward speaking) were assessed before and after treatment, and at a 1- and 2-year follow-up. The results showed that: (a) there was no relationship between stuttering severity and the severity of negative emotions and cognitions, (b) the severe stuttering group had the largest treatment gains but also the highest level of regression, and (c) at post-treatment and both follow-up assessments the differences on measures of emotions between the mild and severe emotional group had disappeared, chiefly due to a large decrease in the latter group's negative emotions and cognitions. Our findings show that, based on treatment gains, specific subgroups can be identified, each requiring different treatment approaches. This underlines the necessity of developing a better understanding of how various dimensions of stuttering relate to treatment outcome.

Introduction

Given the current need to prove effectiveness for all kind of treatments (Sackett, Rosenberg, Gray, Haynes & Richardson, 1996), researchers and clinicians have extensively discussed the designs and methods for efficacy research in stuttering therapy (*e.g.*, Bernstein Ratner, 2005; Bothe, 2003; Conture, 1996; 1997; Conture & Guitar, 1993; Conture & Yaruss, 1993; Cordes & Ingham, 1998; Costello & Ingham, 1984; Curlee, 1993; Finn, 2003a; Ingham, 2003; Ingham & Cordes, 1997; Onslow, 2003; Yaruss, 1998a, 2001). Although many studies have shown that stuttering therapy (*e.g.*, Conture, 1996; Craig, et al., 1996; Craig, 2002; Langevin & Boberg, 1993; Onslow, 2001; Perkins, 2001) is efficacious, not all people who stutter benefit from therapy equally. stuttering therapy is not always effective. The use of fluency enhancing techniques tends to result (at least initially) in unnatural speech (Dayalu & Kalinowski, 2001; Franken, Boves, Peters, & Webster 1995; Ingham, Gow & Costello, 1985a; Ingham, Martin, Haroldson, Onslow & Leney 1985b; Kalinowski, Noble, Armson & Stuart, 1994). It has also become clear that, some individuals are not able to maintain speech gains made in therapy (Boberg, 1981; Bray, Kehle, Lawless & Theodore, 2003; Craig, 1998; Eichstädt, Watt & Girson, 1998; Finn, 2003b; Hasbrouck & Lowry, 1989; Ladoucer & Auger, 1980; Ryan & Van Kirk Ryan, 1995; Wagaman, Miltenberger, & Arndorfer, 1993). However, it remains unclear why some individual benefit more from therapy than others, both in the short and long term. One method to address this issue is to explore whether or not there are different treatment outcomes for clients with differing pre-treatment stuttering profiles.

Many clinicians and researchers have attempted to characterize stuttering behaviors on the basis of various aspects, such as etiology (*e.g.*, Blood, 1985; Poulos & Webster, 1991; St Onge, 1963), recovery (for a review, see Seider, Gladstein & Kidd 1982, 1983) and stuttering characteristics (*e.g.*, Andrew & Harris, 1964; Barber Watson, 1987; Borden, 1990; Schwartz & Conture, 1988; van Riper, 1982; Yaïri, 1990). Borden (1990) reviewed a number of dimensions on which subtypes of stuttering have been differentiated in the past. Her list included among other aspects: severity (mild, moderate, severe); manifestation (covert, overt); locus of stuttering block (labial, laryngeal, respiratory); phonetic features (vowels, consonants); and type of speech behaviors (prolongations, repetitions). She studied which of these classifications are useful to improve our understanding of stuttering and our ability to help those who stutter become more fluent. Borden and others argued that important differences between mild and severe stuttering (Borden, Baer & Kenney, 1985; Watson & Alfonso, 1987) are due to reactive behaviors (secondary factors) and not to stuttering itself (primary factors). How classification or subtyping of stuttering relates to stuttering therapy remains to be investigated.

In the present study, we chose to explore the differences in treatment outcome with reference to two dimensions of stuttering: (1) stuttering severity (primary factors) and (2) the nature and severity of negative emotions and cognitions that are related to stuttering (psychosocial or secondary factors). We included both dimensions because clinical experience suggests that they are related but independent phenomena. Although severity of stuttering tends to correlate with severity of negative introspective clinical characteristics, this is not always the case, as in instances of 'covert' or 'interiorized' stuttering where negative emotional and cognitive reactions are pronounced but overt stuttering is mild. The converse may also occur. Thus, both aspects of stuttering are investigated in this study, which allowed us to analyze the treatment-induced effects for each dimension separately.

Questions concerning the relationship between severity and the extent of treatment gains are being asked with a view to determining whether or not pre-treatment stuttering severity and the severity of negative emotions and cognitions influence response to treatment. A better understanding of how different profiles of stuttering contribute to treatment outcome might help us improve the selection of treatment strategies that address the individual needs of a client (as recommended by McClean, Tasko & Runyan, 2004).

In the current study, the influence of the severity of stuttering and the severity of negative emotions and cognitions on the outcome of therapy is tested on a well-established treatment program that was delivered in the Netherlands: the ISTAR Comprehensive Stuttering Program (CSP; Boberg & Kully, 1985; Kully & Langevin, 1999). The CSP is an integrated therapy that addresses both speech production and related attitudinal problems, and has been shown to produce durable improvement.

In Boberg and Kully (1994), 76% (13 of 17 adults aged 18 to 36 years and 19 of 25 adolescents aged 11 to 17 years) were maintaining satisfactory (≤3% percent syllables stuttered [%SS]) or marginally satisfactory (3.1–6.0 %SS) levels of fluency at 1 and 2 years post-treatment (adults means=1.33 and .96; adolescent means=1.11 and 1.26 %SS). In Langevin and Boberg (1993) 80% of clients who attended a 3-week intensive program for adults (8 of 10 adults aged 16 to 38 years) were maintaining satisfactory or marginally satisfactory levels of fluency at 1 year follow-up (mean=1.3 %SS). In both studies, the outcome measure was 2 minutes of client talk time in telephone calls. At pre-treatment and immediately post-treatment clients made telephone calls to business. At 1 and 2 years follow-up clients received surprise telephone calls made to their home or workplace by research assistants. In addition to improvements in speech, improvements in attitudes, confidence, and perceptions of speech performance were being maintained. Eighty percent of clients in both studies rated their speech at follow-up as satisfactory.

Method

Participants
Twenty-five Dutch adults who stuttered participated in this study (17 men, 8 women; mean age 29.6 years, age range 17-53). Their educational levels ranged from university (n=14; 56%) to pre-university (n=3; 12%), intermediate vocational (n=5; 20%) and lower vocational education (n=3; 12%). All participants were willing and able to attend the three week intensive CSP program which, in this study, was delivered in a residential format. The participants did not attend another treatment program in at least one year before onset of the CSP program. None of the participants had previous experience with the CSP nor with a treatment program similar to the CSP. Prior to treatment onset, all participants were informed of the study, its rationale and content, and signed consents to participate. To be included in the study, clients had to meet the following criteria: (1) a reported onset of stuttering before the age of six; (2) no reported problems in motor development; (3) no reported concurrent problems in speech and/or language development; (4) no reported use of medication that could influence respiration, phonation or articulation; (5) no reported psychiatric problems; and (6) no reported hearing problems.

Stuttering therapy
The CSP is a stuttering treatment that integrates fluency enhancing techniques, tension and stuttering modification techniques, and cognitive behavioral strategies to deal with the emotional and attitudinal aspects of stuttering. A detailed description of the program is presented in Kully and Langevin (1999).

In order to obtain more detailed information on the actual content of the treatments administered, a so-called therapy card[8] (see Appendix A) was designed. After each treatment day, the clinicians filled in the approximate amount of time spent on the main treatment goals and the time dedicated to the strategies or skill-training exercises adopted to achieve these goals (*e.g.*, prolongation, smooth blending, confidence or social skills). The latter aspects were divided into treatment interventions targeting motor control and those directed at emotions and cognitions. The card showed that 73.3% of the therapy time was devoted primarily to skill training exercises targeting speech motor control (*e.g.*, prolongation or smooth blending); 26.7% was devoted primarily to the reduction of the negative emotions and cognitions associated

[8] *The therapy card was jointly designed by a group of seven researchers/clinicians involved in the present effect study and member of the so-called 'advising board'.*

with stuttering. It should be noted that clinicians had to fill in the average time for the whole group of participants. Since many sessions had a dual focus on speech techniques and attitudes, the percentages represent an estimate of the time in which the particular component was the primary focus. Although more time was spent on speech motor control techniques (*e.g.*, on fluency shaping and stuttering modification techniques) the motor control and attitudinal-emotional/cognitive components were (as indicated by the clinicians) equally important at a group level.

To ensure treatment integrity, the CSP program (including both follow up sessions) was delivered in the Netherlands by a team that included two senior ISTAR staff (the second author and ISTAR's clinical coordinator) and a Dutch clinical coordinator who had previously been trained at ISTAR. The team also included eight Dutch trainees, 5 of whom were certified speech, language pathologists in the Netherlands and 3 of whom were students in a speech and language pathology program in The Netherlands. There were no dropouts in the group of participants. At their discretion, clients engaged in a variety of maintenance activities in the post-treatment period, ranging from minimal or no deliberate maintenance activities to continual speech practice and attendance at self-help practice sessions and clinic administered refresher sessions (see Langevin and Kully, 2003).

Instrumentation and procedure

Assessment sessions

Participants were tested immediately before the onset of the therapy[9] (pre-treatment), immediately after the end of the program (post-treatment) and one year (follow-up 1, F1) and two years (follow-up 2, F2) after therapy completion. The participants' performance was assessed at two levels: (1) speech behaviors and (2) emotions and cognitions related to stuttering (see 'Outcome measures' for more details).

In each assessment session, the participants' speech was studied at two levels of observation: (1) s*peech behavior (*stuttering severity, speech quality and speech-motor control) and (2) assessments based on *self-reports*. The whole assessment took approximately two hours. Data with respect to speech motor control (derived from the Nijmegen Speech Motor Test) and data with respect to speech quality (*e.g.*, naturalness judgments) are not further discussed in this paper. To prevent for possible order effects, the order in which different parts of the session were conducted varied in each session. All participants were tested individually in the presence of one experimenter who was not related to the CSP. Four different speech tasks (interview, monologue, reading and telephone call) were recorded on DAT-tape and on video. Data analyzed in this study were taken from the interview samples only because the interview samples most resemble usual talking situations.

[9] *Twelve clients were tested on two (contrary to the one pre therapy session of the remaining clients) different occasions before the start of the therapy (with approximately one month between the two pretreatment measurements). The two datasets were compared with each other to establish variation in speech behaviors across time. The results revealed that there were no significant differences between the two pretreatment assessments (see also Huinck & Peters 2004).*

Independent variables

This study was set up as a factorial design with 'stuttering severity' and 'severity of negative emotions and cognitions' as between-subjects variables, and 'session' (pre, post, F1 and F2) as the within-subject variable.

For the between-subjects variables there were two levels of severity: mild and severe. Although three severity levels (mild, moderate and severe) are often used, we divided the whole group in two, to facilitate the interpretation of the analysis (*e.g.*, the interactions). Based on the severity scores on the stuttering dimension, each participant was classified with either severe stuttering (SS) or mild stuttering (MS). Based on the severity scores on three criterion variables (*i.e.* self-evaluation questionnaires) with respect to negative emotional and cognitive reactions (NECR) related to stuttering, each participant was classified as having either severe negative emotional and cognitive reactions to stuttering (SE) or mild negative emotional and cognitive reactions to stuttering (ME). Table 1 shows individual scores on the criterion variables, ages and educational levels. Groups (see below for a description of the classification) did not differ significantly with respect to age ($F(1,24)=2.371$, $p=0.137$) and educational level ($F(1,24)=0.023$, $p=0.880$).

Classification on the stutter severity dimension

The division of severe and mild on the stuttering dimension was based on the Stuttering Severity Instrument (SSI; Riley, 1980). Participants were allocated to the MS group (n=12) when they had an SSI score $\leq$ 31 (percentile < 77). Five of these participants scored between 21-31 and were, according to the SSI classification norms, moderate stutterers (see Table 1). Participants were allocated to the SS group (n=13) when their SSI score $\geq$ 32 (percentile > 78). The mean SSI scores of the MS group and the SS group respectively were 18.08 (SD=8.62) and 36.31 (SD=3.90).

Classification with respect to 'negative emotional and cognitive reactions' (NECR)

The division into severe and mild negative emotional and cognitive reactions to stuttering was based on the following three self-report instruments which all participants completed before their treatment program started. First, the Avoidance scale of the *Perceptions of Stuttering Inventory (PSI;* Woolf, 1967). This inventory comprises 60 items that represent the parameters of struggle (S), avoidance (A) and expectancy (E) and examines the stutterer's perception of the presence of these factors in his communication process. Second, Lanyon's *Stuttering Severity Scale (SS Scale;* 1967) which is an inventory of stuttering-related feelings, behaviors and attitudes comprising 64 speech situations, which assess the severity of stuttering. The stutterer is asked to indicate whether the true/false statements are applicable or not (*e.g.*, 'I'm sensitive about my speech problems' or 'If I did not stutter, I would probably speak much more than I do'). The scores are compared to a norm score of stuttering speakers (40.6; SD=11.9). And finally, the *Inventory of Interpersonal Situations (IIS)*, a scale developed by Van Dam-Baggen and Kraaimaat, (1987; see also Kraaimaat, Vanryckeghem & van Dam-Baggen, 2002). It measures two components of social anxiety: the extent to which emotional tension or discomfort is perceived in social situations and the frequency with which social responses are executed. Both parts employ the same 35 items to elicit responses to social situations. The 35 statements are grouped in five subscales: 'giving criticism', 'expressing opinion', 'giving a compliment', 'initiating contact', and 'positive self-statements'. Overall, stuttering speakers display significantly higher levels of emotional tension or discomfort in social situations than non-stuttering speakers do.

We selected this particular set of measures for the division into mild and severe, so that different levels of introspective clinical characteristics (*e.g.*, avoidance, stuttering-related feelings, behaviors and attitudes, and social anxiety) were captured. This way, the classification reflects to large extent the actual component of negative emotions and cognitions.

Table 1. Profiles of the participants with respect to Gender (G), age pre treatment (Age), educational level pre treatment (E), and raw scores of the criteria variables: IIS-t and IIS-f (Inventory of Interpersonal Situations, tension and frequency scales), PSI avoidance (PSI-a), Lanyon's stuttering severity scale (Lanyon) and SSI (stuttering severity instrument). Final columns present for each subject the classification into mild and severe stuttering and into mild and severe negative emotions and cognitions (emotional severity).

G	*Age*	*E*[a]	*IIS* *F*	*IIS* *T*	*PSI-a*	*Lanyon*	*Emotional severity*	*SSI*	*Stuttering Severity*
V	18	P	102	85	8	41	Mild	29	Mild
V	49	H	112	102	7	36	Mild	11	Mild
M	22	H	109	85	4	24	Mild	7	Mild
M	30	I	101	83	11	40	Mild	15	Mild
M	36	H	98	98	6	39	Mild	16	Mild
M	22	I	110	69	6	41	Mild	30	Mild
V	28	L	88	100	19	18	Mild	32	Severe
M	18	P	133	77	8	34	Mild	32	Severe
V	17	P	128	74	11	36	Mild	35	Severe
M	23	H	110	89	2	40	Mild	39	Severe
M	17	I	113	80	9	19	Mild	40	Severe
M	32	H	111	83	6	40	Mild	45	Severe
M	28	H	89	106	4	30	Mild	36	Severe
V	24	H	109	106	14	46	Severe	27	Mild
V	48	L	123	84	13	48	Severe	9	Mild
V	53	L	98	88	16	47	Severe	17	Mild
M	19	H	107	86	15	39	Severe	7	Mild
M	49	H	85	106	13	41	Severe	24	Mild
M	28	H	115	98	18	39	Severe	25	Mild
M	26	H	101	117	9	47	Severe	32	Severe
M	23	H	101	114	12	47	Severe	35	Severe
M	24	H	115	69	16	52	Severe	39	Severe
V	37	I	94	107	15	56	Severe	39	Severe
M	23	I	107	108	5	44	Severe	35	Severe
M	47	H	116	77	19	48	Severe	33	Severe

Note. [a]*Educational level is classified as H (higher vocational education and university), P (pre university education), I (Intermediate vocational education), and L (lower vocational education).*

After the scores on each instrument had been ranked from the lowest to the highest level of severity, for each scale subjects were equally divided into three categories: (a) high, (b) intermediate and (c) low. This resulted in three scores of 2, 1, or 0 points, respectively. These points were added to obtain a final score for each participant. Participants were then ranked on the basis of their final scores. Participants in the lower half of the range of final scores were placed in the mild group with respect to negative emotions and cognitions (ME, n=13) and those in the upper half of the range were placed in the severe group (SE, n=12). The mean scores in the ME group were 7.77 (SD=4.28) on the PSIa, 33.69 (SD=8.34) on the Lanyon scale, 108 (SD=13.01) on the IIS-f and 87 (SD=11.41) on the IIS-t. The mean scores in the SE group were 13.75 (SD=3.84) on the PSIa, 46.17 (SD=4.99) on the Lanyon scale, 105.92 (SD=10.71) on the IIS-f and 96.67 (SD=15.45) on the IIS-t. Table 1 shows the individual scores of the participants.

Outcome measures

In the next sections the measures that were used to assess the two dimensions (speech fluency and introspective clinical characteristics) of speech are discussed separately.

Speech fluency measurements
(1) Percentage of stuttered syllables (%SS). Three trained raters (see Boberg & Kully, 1994) for a description of the rater training program) counted the number of stutters and the number of syllables (speech rate) in three-minute speech samples taken from a videotaped within clinic interview, using electronic button-press event recorders (Boberg & Kully, 1985). The speech samples (interview, reading monologue and telephone call) were randomly assigned to the raters, stratified for participant (*i.e.* one rater rated a participant's samples from all four measurement occasions, but did so without identifying information). Eleven percent of the samples were rated by all three raters and then re-rated 6 months later. These samples were randomly selected. Both intra- and inter-rater reliability (inter class-correlation coefficients, Shrout & Fleiss, 1979) for these measurements were calculated. Inter-rater reliability was 0.979 and intra- rater reliabilities were 0.988, 0.994, and 0.985 for each of the three raters respectively.

(2) The Distorted Speech scale of the Speech Situation Checklist (SSC; Brutten, 1975; Brutten & Janssen, 1980). The SSC addresses 51 situations that may arouse negative emotions in persons who stutter (*e.g.*, ordering food in a restaurant, telephone conversation) and consists of two levels or subscales: the distorted speech scale (DS) and the emotional reaction scale (ER; for a description of this latter scale, see the next section). The level of distorted speech is scored on a five-point scale from 1 (not at all) to 5 (very much). The scores can be compared with norm scores (Bakker, 1980) for Dutch stutterers (norm DS =129.6, SD=29.3). It should be noted that the term 'Distorted Speech' refers to 'Speech Disorganization' (Brutten & Janssen, 1980) and to 'Speech Disruption' (Brutten, 1973, 1975).

(3) Diadochokinesis (DDK) task. In this task speakers have to repeat the sequence /pətəkə/ as quickly as possible in 9 seconds. The number of correctly and fluently produced syllables is counted and used as an indication of overall speech motor production capacity. This measure is related to the degree of control the speaker has over his speech system and is therefore taken to indicate stuttering severity (see also Huinck, Wouters, Hulstijn & Peters, 2001).

Measures of introspective clinical characteristics
(1) The Emotional Reaction scale of the Speech Situation Checklist (SSC; Brutten, 1975; Brutten & Janssen, 1980). The ER subscale of the SSC (complete scale is described above) measures the level of emotional reaction on a five-point scale from 1 (not at all) to 5 (very much). The scores can be compared with norm scores (Bakker, 1980) for Dutch stutterers (norm ER=129.3, SD=30.5).

(2) Speech satisfaction rating scale. The participants judged their overall speech satisfaction on an eleven-point self-rating scale, with 0 as the worst possible and 10 as the best-possible judgment. This is parallel to the grading system for academic performance used in Dutch schools where 1 to 5 denotes insufficient, and 6 to 10 sufficient performance.

(3) The S24 attitude scale (Andrews & Cutler, 1974; Erickson, 1969). This scale deals with the speaker's attitudes toward his communication ability. Higher scores reflect more negative attitudes.

Statistical analysis

Two multivariate mixed-design analyses of variance with repeated measures were performed: one for the speech measures and one for the measures related to negative introspective clinical characteristics. Fifteen missing values (of the 600 data points) were imputed with estimates calculated with the SPSS 12.01 regression-based missing value analysis. To determine the dimension(s) on which the mild and severe groups showed different treatment results, in each analysis we used one within-subject factor, *i.e. assessment 'session'* (pre-treatment, post-treatment, Follow-up 1 and Follow-up 2) and two between-subject factors (dimensions), *i.e. stuttering severity* (mild versus severe stuttering) and emotional severity (mild versus severe negative emotions and cognitions). Because Maughly's Test of Sphericity showed a high epsilon on all measures, averaged multivariate results (p-values) are given (Huynh-Feldt corrected). Univariate results are reported to illustrate the differences between the groups on all measures. Within-subject contrasts (Bonferroni corrected for outcome measure) are used to show the significance of the differences between pre- and post-treatment, between pre-treatment and F1 and between pre-treatment and F2.

Furthermore, as an indication of the external validity of the classification system, Pearson correlations were performed between the classification measures and the pre-treatment dependent variables.

Results

Table 2 summarizes the multivariate results, organized in measures of speech and measures of introspective clinical characteristics (ICC). Because there was no significant three-way interaction, the resultant p- and F-values were not included in this Table.

In Table 3, F-values, mean differences and significance levels of the univariate results are presented.

*Table 2. Results of the multivariate tests expressed in F-values and degrees of freedom (df). Significance is indicated by * (p<0.05) and ** (p<0.01).*

(Wilks' Lambda)	Between subjects			Within subjects		
	SS[a]	NECR[b]	SS x NECR	Session	Session x SS	Session x NECR
	F(df)	F(df)	F(df)	F(df)	F(df)	F(df)
Speech measures	9.17** (3,19)	0.84 (3,19)	2.51 (3,19)	12.38** (9,149)	2.81** (9,149)	2.29* (9,149)
Measures of ICC[c]	0.70 (3,19)	0.42 (3,19)	1.18 (3,19)	16.29** (9,149)	2.75** (9,149)	2.35* (9,149)

note. [a]SS denotes mild versus severe with respect to 'Stuttering Severity', [b]NECR denotes mild versus severe with respect to 'negative emotional and cognitive reactions', [c]ICC denotes introspective clinical characteristics.

(1) Speech measures

Multivariate results

Results showed a significant two-way interaction between 'session' and 'emotional severity' (ME versus SE), showing that the overall differences between the severe and the mild group varied over the four assessments (see Table 2 for F-values, degrees of freedom and significance levels).

The interaction between 'session' and 'stuttering severity' (MS versus SS) was also significant, indicating an effect of stuttering severity on the speech-related treatment results. Furthermore, there was a significant main effect of 'session' (pre, post, F1 and F2) and of 'stuttering severity'. Obviously, these main effects should be interpreted in the context of the two-way interaction effects described above. There was no main effect of 'emotional severity' and, with an F-value below 1, there was not even a trend. To facilitate a lucid interpretation of the multivariate results, we will describe the findings for the univariate tests next.

Table 3 Univariate results of the speech measures and the measures of introspective clinical characteristics (ICC).

Speech measures	*%SS*			*Brutten DS*			*DDK*		
	$F^{b,d}$	*Mean difference*	*SE*	$F^{b,d}$	*Mean difference*	*SE*	$F^{b,d}$	*Mean difference*	*SE*
Session[a]	14.96****	10.25	1.57	42.47****	150.71	4.66	8.62****	13.38	0.96
Pre-post[c]	30.68****	9.17	1.655	86.14****	56.41	6.08	12.87***	-2.91	0.81
Pre-F1[c]	11.42***	3.09	0.913	37.27****	24.99	4.09	24.23****	-3.45	0.70
Pre-F2[c]	19.21****	3.79	0.866	41.33****	27.51	4.28	8.40***	-2.63	0.91
NECR (ME-SE)	2.65	-7.78	2.442	0.02	-16.11	7.18	0.30	3.06	1.25
SS (MS-SS)	10.14***	3.98	2.442	5.03*	1.12	7.18	5.97*	0.68	1.25
NECR*session	1.25			3.53*			1.77		
SS*session	5.20***			1.20			1.12		
ICC measures	*Brutten ER*			*Speech satisfaction*			*S24*		
	$F^{b,d}$	*Mean difference*	*SE*	$F^{b,d}$	*Mean difference*	*SE*	$F^{b,d}$	*Mean difference*	*SE*
Session[a]	31.89****	144.85	3.97	36.38****	4.82	0.27	22.19****	16.10	0.75
Pre-post[c]	99.63****	43.10	4.318	96.20****	-3.22	0.33	41.43****	7.20	1.12
Pre-F1[c]	15.57***	19.32	4.896	18.99****	-1.56	0.36	28.43****	4.55	0.85
Pre-F2[c]	31.05****	21.85	3.924	50.81****	-1.72	0.24	53.89****	5.65	0.77
NECR (ME-SE)	0.14	-2.42	6.590	0.005	-0.02	0.28	0.919	-1.19	1.25
SS (MS-SS)	0.14	-2.46	6.590	1.48	0.34	0.28	0.011	-0.13	1.25
NECR*session	6.28***			0.61			2.85*		
SS*session	3.27*			5.36**			2.25		

Note. [a]*The Main factor 'session' is Huynh-Feldt corrected. In this row, under 'Mean difference' the pre-treatment mean and SE is given.*
[b]*Degrees of freedom were 3,63 for the factor 'session' and for the interactions NECR (negative emotional and cognitive reactions) severity *'session' and stuttering severity (SS)*'session'. Degrees of freedom were 1,21 for the contrasts and the between subject (ME-SE and MS-SS) effects.* [c]*Within subjects contrasts* [d]$p<0.0001$****; $p<0.001$***; $p<0.01$**; $p<0.05$*

Univariate results

Results showed (see also Table 3): (1) a significant overall main effect of therapy on all three measures of speech (%SS, Brutten DS and DDK); (2) significant contrasts between pre- and post-treatment at all three measures of speech, showing improvements relative to pre therapy; (3) that even though the mean scores of all speech measures showed a clear regression at both follow-ups (F1 and F2), significant gains relative to pre-treatment levels, were maintained (see Figures 1a and 2a); (4) that there were no differences between the ME group and the SE group in the speech measures, and no significant interactions between 'session' and 'emotional severity'. In F1 and F2 both groups exhibited a regression toward pre-treatment levels (see Figure 2a); (5) a significant difference between the MS and the SS group on all three measures of speech, but only on the %SS was there a significant interaction between 'session' and 'stuttering severity'. The SS group showed a higher %SS and more post-treatment absolute gains but also more regression in F1 and F2 relative to the MS group at all measurement sessions.

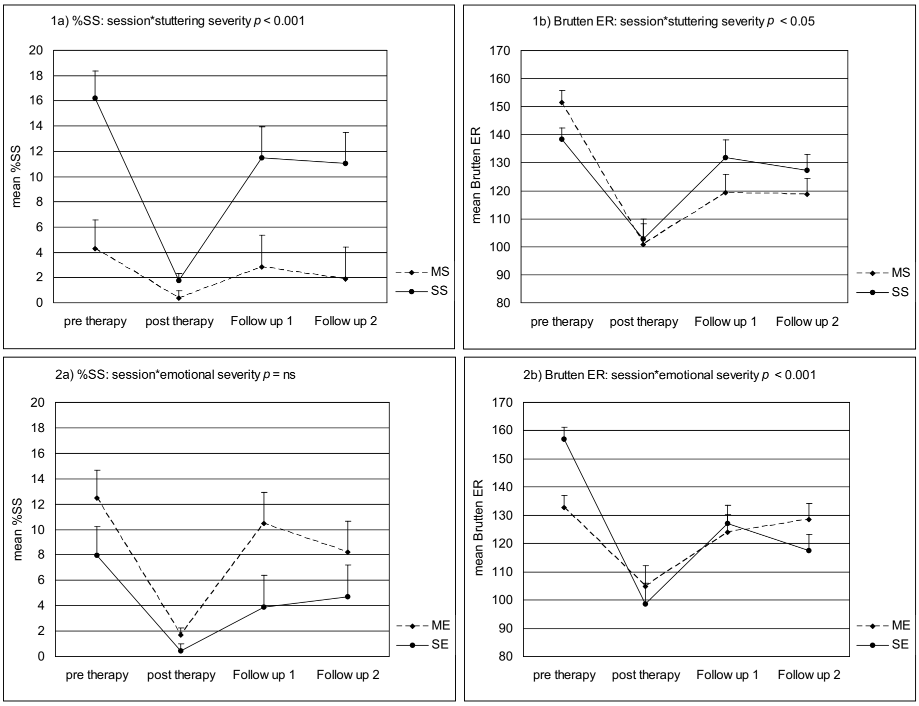


Figure 1. The mean scores (and standard errors) for the MS and SS group on the %SS (1a) and on the Brutten ER-scale (1b) and for the ME and SE group on the %SS and Brutten ER scale (2b). P-values (ns=not significant) are given at the top of each figure.

(2) Measures of introspective clinical characteristics
Multivariate results
There was a significant two-way interaction between 'session' and 'negative emotional and cognitive reactions' (ME versus SE), showing that the overall differences between the severe and mild group varied over the four assessment sessions (see Table 2 for F-values, degrees of freedom and significance levels).

As in the speech measures, there was a main effect for 'session'. However, the effect was again qualified by the interaction between 'session' and 'stuttering severity' (MS versus SS), implying that stuttering severity influenced the change in emotions and cognitions over the course of therapy and the two follow-ups. The main effect for 'stuttering severity' and for 'severity of negative emotional and cognitive reactions' was not significant. Moreover, both F-values were below 1, indicating that there was not even a trend.

Univariate results
Results showed (see also Table 3): (1) significant positive main effects of session on all three measures of introspective clinical characteristics; (2) significant pre-post-treatment contrasts on all three measures; (3) that the significant gains were maintained at both follow-ups (F1 and F2), even though the mean scores showed a clear regression towards pre-treatment levels at F1 and F2; (4) no overall difference between the ME and SE groups but a significant interaction between 'session' and 'emotional severity' on the Brutten ER (Figure 2b) and the S24; (5) that the MS-SS comparison also failed to yield a group difference in the measures of introspective clinical characteristics, but again, that the interaction between 'session' and 'stuttering severity' was significant on the Brutten ER and speech satisfaction rating. Remarkably, on the Brutten ER the MS-group showed a higher mean pre-treatment score, larger post-treatment gains, and less regression in F1 and F2 when compared to the SS group. However, the speech satisfaction score of the MS group was higher (indicating a more positive speech satisfaction) than the score of the SS group at the pre treatment measurement. The difference disappeared post treatment. The MS group showed regression at F1 but the score increased again at F2. The SS group showed also little regression at F1 but the speech satisfaction score decreased further at F2.

Pearson correlations
Pearson correlations between the criterion variables for rating stuttering severity (SSI) and the dependent pre-treatment speech measures (%SS, Brutten DS and DDK) revealed a significant ($p<0.01$) correlation ($r=0.555$) between the SSI score and the pre-treatment %SS. The pre-treatment scores on the DDK and Brutten DS did not correlate.

Pearson correlations between the criterion variables for classifying into mild and severe negative emotions and cognitions and the dependent measures of negative emotions and cognitions demonstrated a significant ($p<0.01$) correlation ($r=0.482$) between Lanyon's SS scale and Brutten's ER and a significant ($p<0.01$) correlation ($r=0.759$) between Lanyon's scale and the S24.

Pearson correlations between the two types of criterion variables (ISS and negative emotional and cognitive reactions) showed no relationship ($r=-0.061$) between these classification systems.

Discussion and conclusions

In the present study we explored whether we would find differences in the treatment outcomes of specific subgroups of stuttering individuals. Based on the pre-treatment scores on self-report questionnaires and the stuttering severity instrument (SSI, Riley, 1980), the participants were characterized as having either mild or severe negative emotions and cognitions (emotional severity, ME versus SE) associated with their stuttering and as having either a mild or a severe stuttering problem (stuttering severity, MS versus SS). For each dimension, the short- and long-term treatment results of the groups classified as severe, were compared to those of the groups that were classified as mild.

The main findings were: (1) Different stuttering profiles show different treatment outcomes. (2) The largest treatment gains (in terms of absolute gain) but also greatest regression occurred in the severe stuttering (SS) group. (3) The absent relationship between stuttering severity and the severity of negative emotional and cognitive reactions. Below, these findings are discussed in more detail, followed by a discussion of the limitations of this study and some concluding remarks.

(1) Different stuttering profiles show different treatment outcomes.
Effects of stuttering severity (MS versus SS) on speech measures.
Our multivariate tests revealed clear differences between the mildly (MS) and severely stuttering (SS) participants. These groups differences over time were mainly caused by one of the three variables: the percentage of stuttered syllables (%SS). The difference itself was not surprising because the group classification method was based on the SSI scores and these correlated significantly with the %SS. However, immediately after therapy the dissimilarity had been greatly reduced. The SS group had gained much more in fluency (in terms of absolute gain) than the MS group. At the two follow-ups both groups showed regression but their respective proportions remained below pre-treatment levels. Moreover, at F1 and F2 the initial differences between the two groups had reappeared.

As much as the mild and severe groups differed with respect to pre-treatment dysfluency (%SS), no differences were found in the pre-treatment scores of the Brutten distorted speech (DS) scale. Obviously, the self-reported pre-therapy level of perceived dysfluency (specifically that of the mild group) did not fully correspond to the actual level of dysfluency.

Effects of stuttering severity (MS versus SS) on measures of introspective clinical characteristics.
There was no direct link between stuttering severity and negative emotions and cognitions. This was particularly seen in the Brutten ER and speech satisfaction scores. Relative to the SS group, the pre-treatment data showed a higher mean ER for the MS group but at post-treatment assessment the group scored similarly, indicating larger treatment effects in the MS group. This suggested an absent relation between stuttering severity and the severity of negative emotions and cognitions. Figure 1b suggests that at the long-term the mild stutterers maintained most gains: they now had even lower mean ER scores than the severe stutterers.

With respect to self-reported speech satisfaction, the pre-treatment assessment showed that the SS group had assigned themselves the lowest ratings but that this group difference disappeared following therapy. The post-treatment scores for both groups had increased to around 8, reflecting a high level of speech satisfaction. Despite their relative regression in the longer term, both groups were still satisfied with their speech one and two years after therapy.

It is plausible that the stuttering therapy (CSP) not only improved the participants' fluency but also their emotions and cognitions, specifically in the mild stuttering group. After all, 26% of the CSP treatment time was devoted to cognitive behavioral aspects of therapy. The fewer the dysfluencies they experienced, the fewer associated negative emotions they reported immediately following the end of treatment. It is similarly plausible that after such greatly enhanced fluency these negative emotions would have returned when they failed to maintain the initial level of fluency during the long-term follow-up period. This was particularly true for the severe stutterers. And yet, disfluency and negative reactions to speech did not affect their speech satisfaction ratings. Both groups remained satisfied in the long term (ratings > 5). Despite the low pretreatment %SS of the MS group (Figure 1a), higher emotional reactions (Brutten ER) were experienced in this group when compared to the SS group (Figure 1b). The latter group demonstrated a much higher %SS pretreatment but scored below the MS group on this self-assessment scale. This suggests that individuals with mild stuttering profiles may profit most from interventions that help them put their stuttering into perspective, thus decreasing their negative emotions and cognitions, whereas the treatment of speakers with a severe profile may first need to focus on reducing the stuttering. Relapse in %SS on the other hand seemed to affect the emotional reactions importantly. The substantial relapse in the SS group coincided with regression in negative emotional reactions that were even greater than those of the MS group. The relapse experienced by the sever group suggests that they may require additional therapy to facilitate maintenance of speech and cognitive emotional gains.

Effects of emotional severity (ME versus SE) on speech measures.

There was no overall (multivariate) difference between the mild (ME) and severe emotional (SE) group on the speech measures but the interaction between the effect of therapy and 'group' was significant. For an adequate interpretation of the data of the three fluency measures, it should be noted that, perhaps contrary to expectations, the SE group had a lower mean %SS than the ME group at each assessment. Thus, even though they had very severe attitudinal problems, their speech was relatively fluent throughout (as reflected by the absence of a significant interaction between 'group' and 'session'). This result challenges Borden's (1990) suggestion that some of the key differences between mild and severe stuttering (see also Borden, Baer & Kenney, 1985; Watson & Alfonso, 1987) are due to reactive behaviors (secondary factors) and not to stuttering itself (primary factors). The pre-treatment distorted speech scores (Brutten DS) present a different picture: the SE group scores exceeded those of the ME group. This indicates that, despite their relatively low proportion of stuttered syllables, the high-emotional stutterers judged their own speech as being more distorted than the mild-emotional stutterers. At all three post-treatment tests the groups reported similar DS scores.

Effects of emotional severity (ME versus SE) on measures of introspective clinical characteristics.
We did not find an overall (multivariate) difference between the mild and severe emotional groups on the emotional measures. We expected emotional severity to be reflected in the psychosocial measures (as stuttering severity would be by the speech fluency measures). Only the Brutten ER and the S24-attitude scale yielded differences between the severe and mild emotional groups and, moreover, only at pre-treatment assessment. After therapy most of these differences had disappeared, which was basically attributable to the large decrease in negative emotions and cognitions in the SE group.

(2) The largest treatment gains but also greatest regression occurred in the severe stuttering (SS) group.
The relatively high levels of relapse in our group of severe stutterers are in line with the levels reported in previous research investigating correlations between %SS and relapse (*e.g.*, Craig, 1996; Guitar, 1976). Craig also reported that high levels of pre-treatment %SS were significantly related to long-term (one year post-treatment) relapse risks. In contrast, Ladouceur and colleagues (1989) found that a high pre-treatment stuttering level was associated with a lack of improvement, whereas the mildly and moderately stuttering participants had clinically improved after treatment and at a six-month follow-up. Contrary to Ladouceur et al., we found that the severe stuttering group made marked improvements after therapy. However, consistent with Guitar and Craig's findings, this group also showed substantial regression toward pre-treatment levels. It should be noted that these are group findings and that, as such, they do not necessarily represent the performance of individuals who stutter severely pre-treatment.

(3) The absent relationship between stuttering severity and the emotional severity (negative emotional and cognitive reactions)
Although initially not one of our research questions, this finding needs further explanation. It relates to the classification system we used to distinguish the participants' profiles of stuttering. The data presented in Table 1 showed a relatively equal distribution of subjects in each cell (SE/SS n=6; ME/SS n=7; SE/MS n=6 and ME/MS n=6), indicating that the level of emotional severity (based on pre-treatment levels of avoidance, experiences of stuttering severity, social anxiety and achievement motivation) was not related to the level of stuttering severity (based on the SSI). This result supports the basic idea that the two dimensions (stuttering severity and emotional severity) need to be investigated separately, as was done in the present study. However, it does not mean that the two dimensions never affect each other. In her 1997 review, Lewis did find evidence for a correlation between pre-treatment communication attitudes, as assessed by four different self-report questionnaires, and stuttering severity, as based on self-rated severity scores, %SS or %words stuttered and a derived score combining stuttering frequency and speech rate. The discrepancy between Lewis' findings and ours may be due to a difference in definition. In our study, speech satisfaction scores were classified as a measure of introspective clinical characteristics whereas Lewis took them to indicate stuttering severity. It is likely that speech satisfaction scores provide information on both aspects of stuttering. By investigating them separately, as was done in the current study, we will develop a better understanding of how treatment affects these two dimensions of stuttering.

Furthermore, only one of the scales used for the a-priori classification with respect to negative introspective clinical characteristics, *i.e.* Lanyon's stuttering severity scale, correlated with the outcome measures for emotional involvement (Brutten ER and the attitude scale S24 tested at baseline). A high correlation between classification measures and outcome measures typically implies that the classification system has a high external validity. On the other hand, a low correlation signifies that different, supplementary aspects of speech were measured. This is also shown by the deviating score pattern of the speech satisfaction rating scale. This scale measured a different aspect than the Brutten ER and the S24. In fact, the speech satisfaction list did not differentiate between mild and severe negative introspective clinical characteristics

Limitations of the study

Although dividing the participants into two groups (severe and mild) facilitated the interpretation of the results and was necessary to increase statistical power, this dichotomization was also a limitation of the study. There are two reasons for this. First, stuttering severity is not bipolar but rather a continuum and second, it is not a stable trait (See Conture, 2001; Crowe, DiLollo, & Crowe, 2000; Manning, 2001). There are of course many stuttering participants who where moderate severe and who were by this division forced to either side of the classification. These subjects might have reduced the group differences, as a result of regression to the mean. In future research this problem should be addressed.

Of course, as all clients followed the same treatment, this study does not particularly indicate how treatment can be adapted to these pre-existing clinical profiles, nor do we know how these profiles interact with a range of treatment options. These are areas for future research.

In conclusion, the present study showed that individuals with different profiles of stuttering respond differently to stuttering therapy. Results suggested that individuals with mild stuttering profiles may profit most from interventions that help them put their disfluency into perspective, thus decreasing their negative emotions and cognitions, whereas the treatment of speakers with a severe profile may first need to focus on reducing the disfluency. Furthermore, no relationship between stuttering severity and the severity of negative emotions and cognitions associated with the disfluency was found. Future research could therefore adhere to an independent investigation of these two distinct dimensions of stuttering. In addition, research efforts should be aimed at the further development of a simple, reliable screening tool to facilitate an adequate classification of stuttering subtypes.

The fact that in our investigations treatment outcome proved to be associated with the clients' severity of stuttering -with the largest gains and also the largest regression in the severely stuttering participants- suggests that, after therapy completion, severe stutterers may need extra attention and coaching in the follow-up period. This does not seem to apply to individuals that have severe emotional problems associated with their disfluency. The current findings suggest that with therapy this type of speaker is capable of improving his/her attitude toward stuttering by forming more realistic beliefs about their speech.

The next step is to develop new approaches or fine-tune existing therapies to fit the client's stuttering profile. In-depth knowledge about the way the various dimensions of stuttering contribute to treatment outcome will help us optimize the efficacy of future stuttering programs.

Acknowledgements

This research project was funded by the Netherlands Health Care Insurance Board (CVZ, College voor Zorgverzekeringen). The authors are grateful to Toni Rietveld and Rogier Donders for their statistical advice and to Pascal van Lieshout for his valuable comments on prior versions of this manuscript. We thank Esther Wouters for her assistance in scoring the questionnaires.

References

Andrews, G., & Cutler, J. (1974). Stuttering therapy: The relation between changes in symptom level and attitudes. Journal of Speech and Hearing Disorders. *Journal of Speech and Hearing Disorders, 39,* 312-319.

Andrews, G., & Harris, M. (1964). *The syndrome of stuttering*. London: Heinemann.

Ansel, B.M. (1993). Treatment efficacy research in stuttering. *Journal of Fluency Disorders, 18,* 121-123.

Bakker, K. (1980). Normering 'vragenlijst voor spreeksituaties' (Brutten) [Dutch standardization of the Speech Situations Checklist]. *Logopedie en Foniatrie 52*, 520-526.

Barber Watson, J. (1987). Profiles of stutterers' and nonstutterers' affective, cognitive, and behavioral communication attitudes. *Journal of Fluency Disorders, 12,* 389-405.

Bernstein Ratner, N. (2005). Evidence-based practice in stuttering: Some questions to consider. *Journal of Fluency Disorders, 30,* 163-188.

Blood, G.W. (1985). Laterality differences in child stuttering: Heterogeneity, severity levels, and statistical treatments. *Journal of Speech and Hearing Disorders, 50,* 66-72.

Boberg, E. (1981). Maintenance of fluency: An experimental program. In E. Boberg (Ed.), New York: Elsevier North-Holland.

Boberg, E., & Kully, D. (1985). *Comprehensive stuttering program*. San Diego, CA: College-Hill Press.

Boberg, E., & Kully, D. (1994). Long-term results of an intensive treatment program for adults and adolescents who stutter. *Journal of Speech and Hearing Research, 37,* 1050-1059.

Borden, G.L., Baer, T., & Kenney, M.K. (1985). Onset of voicing in stuttered and fluent utterances. *Journal of Speech and Hearing Research, 28,* 363-372.

Borden, G.L. (1990). Subtyping adult stutterers for research purposes. *ASHA reports, American Speech Language Hearing Association, 18,* 58-62.

Bothe, A.K. (2003). Evidence-based treatment of stuttering: V. The art of clinical practice and the future of clinical research. *Journal of Fluency Disorders, 28,* 247-258.

Bray, M.A., Kehle, T.J., Lawless, K.A., & Theodore, L.A. (2003). The relationship of self-efficacy and depression to stuttering. *American Journal of Speech-Language Pathology, 12,* 425-431.

Brutten, G.J. (1975). Stuttering: Topography, assessment and behavior change strategies. In J.Eisenson (Ed.), *Stuttering: A second symposium.* New York: Harper & Row.

Brutten, G.J., & Janssen, P. (1980). A normative and factor analysis study of the responses of Dutch and American stutterers to the Speech Situations Checklist. Paper presented at the meeting of the 18th International Association of Logopedics and Phoniatrics conference, Washington, DC.

Conture, E.G. (1996). Treatment efficacy: Stuttering. *Journal of Speech and Hearing Research, 39,* 518-526.
Conture, E.G. (1997). Evaluating childhood stuttering. In R.F. Curlee & G.M. Siegel (Eds.), *Nature and treatment of stuttering: new directions* (pp. 239-256). Needham Heights, MA: Allyn & Bacon.
Conture, E.G. (2001). *Stuttering: Its nature, diagnosis, and treatment.* Boston: Allyn & Bacon.
Conture, E.G., & Guitar, B. (1993). Evaluating efficacy of treatment of stuttering. *Journal of Fluency Disorders, 18,* 253-287.
Conture, E.G., & Yaruss, J.S. (1993). *Handbook for childhood stuttering: a training manual.* Tucson, AZ: Bahill Intelligent Computer Systems.
Cordes, A.K., & Ingham, R.J. (1998). *Treatment efficacy for stuttering.* San Diego: Singular Publishing Group, Inc.
Costello, J.M., & Ingham, J.C. (1984). Assessment strategies for stuttering. In R.F. Curlee & W.H. Perkins (Eds.), *Nature and treatment of stuttering: new directions* (pp. 303-334). San Diego: College-Hill Press.
Craig, A. (1998). Relapse following treatment for stuttering: a critical review and correlative data. *Journal of Fluency Disorders, 23,* 1-30.
Craig, A. (2002). Fluency outcomes following treatment for those who stutter. *Perceptual and Motor Skills, 94,* 772-774.
Craig, A., Hancock, K., Chang, E., McCready, C., Shepley, A., McCaul, A., Costello, D., Harding, S., Kehran, R., Masel, C., & Reilly, K. (1996). A controlled trial for stuttering in persons aged 9 to 14 years. *Journal of Speech and Hearing Research 39*, 808-826.
Crowe, T.A., DiLollo, A., & Crowe, B.T. (2000). *Crowe's Protocols.* San Antonio, TX: Psychological Corp.
Curlee, R.F. (1993). Evaluating treatment efficacy for adults: Assessment of stuttering disability. *Journal of Fluency Disorders, 18,* 151-158.
Dayalu, V.N., & Kalinowski, J. (2001). Re: Stuttering therapy results in pseudofluency. *International Journal of Language & Communication Disorders, 36,* 405-408.
Eichstaedt, A., Watt, N., & Girson, J. (1998). Evaluation of the efficacy of a stutter modification program with particular reference to two new measures of secondary behaviors and control of stuttering. *Journal of Fluency Disorders, 23,* 231-246.
Erikson, C.W. (1969). Assessing communication attitudes among stutterers. *Journal of Speech and Hearing Research, 12,* 711-724.
Finn, P. (2003a). Evidence-based treatment of stuttering: II. Clinical significance of behavioral stuttering treatments. *Journal of Fluency Disorders, 28,* 209-218.
Finn, P. (2003b). Addressing generalization and maintenance of stuttering treatment in the schools: a critical look. *Journal of Communication Disorders, 36,* 153-164.
Franken, M.C., Boves, L., Peters, H.F.M., & Webster, R.L. (1995). Perceptual rating instrument for speech evaluation of stuttering treatment. *Journal of Speech and Hearing Research, 38,* 280-288.
Frattali, C.M. (1998a). *Measuring outcomes in speech-language pathology.* New York: Thieme.
Frattali, C.M. (1998b). Outcomes measurement: Definitions, dimensions, and perspectives. In C.M.Frattali (Ed.), *Measuring outcomes in speech-language pathology.* New York: Thieme.
Guitar, B. (1976) Pretreatment factors associated with the outcome of stuttering therapy. *Journal of Speech and Hearing Research 19*, 590-600.
Hasbrouck, J.M., & Lowry, F. (1989). Elimination of stuttering and maintenance of fluency by means of airflow, tension reduction and discriminative stimulus control procedures. *Journal of Fluency Disorders, 14,* 165-183.
Huinck, W.J., Wouters, E.H.A., Hulstijn, W., & Peters, H.F.M. (2001). Differences between stuttering and non-stuttering people: Comparing diadochokinesis, sentences repetition and reaction time tasks. In B. Maassen, W. Hulstijn, R.D. Kent, H.F.M. Peters, &

P.H.H.M. van Lieshout (Eds.), *Speech motor control in normal and disordered speech* (pp. 311-315). Nijmegen, the Netherlands: Vantilt.

Huinck, W.J., & Peters, H.F.M. (2004). Effetiviteitonderzoek bij stottertherapieen. Een observationale studie naar therapeutische effecten van drie verschillende therapievormen. Unpublished manuscript, Radboud University Nijmegen.

Ingham, J.C., (2003). Evidence-based treatment of stuttering: I. Definition and application. *Journal of Fluency Disorders, 28,* 197-207.

Ingham, J.C., & Cordes, A.K. (1997). Self-measurement and evaluating stuttering treatment efficacy. In *Nature and treatment of stuttering: new directions* (2nd ed., pp. 413-437).

Ingham, R.J., & Bothe, A.K. (2002). Letter to the editor: Thomas and Howell (2001): Yet another "exercise in mega-silliness". *Journal of Fluency Disorders, 27,* 169-174.

Ingham, R.J., Gow, M., & Costello, J.M. (1985). Stuttering and speech naturalness: Some additional data. *Journal of Speech and Hearing Disorders, 50,* 217-219.

Ingham, R.J., Martin, R.R., Haroldson, S.K., Onslow, M., & Leney, M. (1985). Modification of listener-judged naturalness in the speech of stutterers. *Journal of Speech and Hearing Research, 28,* 495-504.

Kalinowski, J., Noble N., Armson, J., & Stuart, A. (1994). Naturalness ratings of the pretreatment and post-treatment speech of adults with mild and severe stuttering. *American Journal of Speech-Language Pathology, 3,* 61-66.

Kraaimaat, F., Vanryckeghem, M., & Van Dam_Baggen, R. (2002). Stuttering and social anxiety. *Journal of Fluency Disorders, 27,* 319-331.

Kully, D., & Langevin, M. (1999). Intensive treatment for stuttering adolescents. In R.F.Curlee (Ed.), *Stuttering and related disorders of fluency* (2nd ed., pp. 139-159). New York: Thieme Medical Publishers.

Ladouceur, R., & Auger, J. (1980). Where have all the follow ups gone? *Behavior Therapy, 3,* 10-11.

Ladouceur, R., Caron, C., & Caron, G. (1989). Stuttering severity and treatment outcome. *Journal of Behavior Therapy and Experimental Psychiatry, 20,* 49-56.

Langevin, M., & Boberg, E. (1993). Results of an intensive stuttering therapy program. *JSLPA, 17,* 158-166.

Langevin, M., & Kully, D. (2003). Evidence-based treatment of stuttering: III. Evidence based practice in a clinical setting. *Journal of Fluency Disorders, 28,* 219-236.

Lanyon, R.I. (1967). The measurement of stutter severity. *Journal of Speech and Hearing Research, 10,* 836-843.

Lewis, K.E. (1997). Research on communication attitude and stuttering in adults: a critical analysis. *Journal of Developmental and Physical Disabilities, 9,* 47-58.

Manning, W.H. (2001). *Clinical Decision Making in Fluency Disorders (2nd Edition).* Vancouver, Canada: Singular Publishers.

McClean, M., Tasko, S. M., & Runyan, C. M. (2004). Orofacial movements associated with fluent speech in persons who stutter. *Journal of Speech, Language, and Hearing Research, 47,* 294-303.

Neumann, K., Euler, H.A., von Gudenberg, W.A., Giraud, A., Lanfermann, H., Gall, V., & Preibisch, C. (2003). The nature and treatment of stuttering as revealed by fMRI. A within- and between-group comparison. *Journal of Fluency Disorders, 28,* 381-410.

Onslow, M. (2001). Re: Frequency altered feedback as an alternative to 'prolonged speech' techniques for the control of stuttered speech. *International Journal of Language & Communication Disorders, 36,* 409-411.

Onslow, M. (2003). Evidence-based treatment of stuttering: IV. Empowerment through evidence-based treatment practices. *Journal of Fluency Disorders, 28,* 237-245.

Perkins, W.H. (2001). Re: A commentary on pseudofluency. *International Journal of Language & Communication Disorders, 36,* 413-415.

Poulos, M.G., & Webster, W.G. (1991). Family history as a basis for subgrouping people who stutter. *Journal of Speech and Hearing Research, 34,* 5-10.
Ryan, B.P., & Van Kirk Ryan, B. (1995). Programmed stuttering treatment for children: comparison of two establishment programs through transfer, maintenance and follow up. *Journal of Speech and Hearing Research, 38,* 75.
Riley, G.D. (1980). *Stuttering severity instrument.* Tigard: C.C. Publications.
Sacket, D.L., Rosenberg, W.M.C., Gray, J.A.M., Haynes, R.B., & Richardson, W.S. (1996). Evidence based medicine: what it is and what it isn't. *British medical journal.* Retrieved 01-13-1996, from http://cebm.jr2.ox.ac.uk/ebmisisnt.htm
Schwartz, H.D., & Conture, E.G. (1988). Subgrouping young stutterers: preliminary behavioral observations. *Journal of Speech and Hearing Research, 31,* 62-71.
Seider, R.A., Gladstien, K.L., & Kidd, K.K. (1982). Language onset and concomitant speech and language problems in subgroups of stutterers and their siblings. *Journal of Speech and Hearing Research, 25,* 482-486.
Seider, R.A., Gladstien, K.L., & Kidd, K.K. (1983). Recovery and persistence of stuttering among relatives of stutterers. *Journal of Speech and Hearing Disorders, 48,* 402-409.
St.Onge, K. (1963). The stuttering syndrome. *Journal of Speech and Hearing Research, 6,* 279-289.
Starkweather, C.W. (1993). Issues in the efficacy of treatment for fluency disorders. *Journal of Fluency Disorders, 18,* 151-168.
Thomas, C., & Howell, P. (2001). Assessing efficacy of stuttering treatments. *Journal of Fluency Disorders, 26,* 311-333.
Van Dam-Baggen, C.M.J., & Kraaimaat, F.W. (1987). *Inventarisatielijst Omgaan met Anderen.* Lisse: Swets & Zeitlinger b.v.
Van Riper, C. (1982). *The Nature of Stuttering.* (2nd ed.) Englewood Cliffs NJ: Prentice Hall.
Wagaman, J.R., Miltenberger, R.G., & Arndorfer, R.E. (1993). Analysis of a symplified treatment for stuttering in children. *Journal of applied behavior analysis, 26,* 53-61.
Watson, B.C., & Alfonso, P.J. (1987). Physiological bases of acoustic LRT in nonstutterers, mild stutterers, and severe stutterers. *Journal of Speech and Hearing Research, 30,* 434-447.
Woolf, G. (1967). The assessment of stuttering as struggle, avoidance and expectancy. *British Journal of Disorders of Communication, 2,* 158-171.
Yairi, E. (1990). Subtyping child stutterers for research purposes. *ASHA reports, American Speech Language Hearing Association, 18,* 50-57.
Yaruss, J.S. (1998). Describing the consequences of Disorders: Stuttering and the International Classification of Impairments, Disabilities, and Handicaps. *Journal of Speech, Language, and Hearing Research, 41,* 249-257.
Yaruss, J.S. (2001). Evaluating treatment outcomes for adults who stutter. *Journal of Fluency Disorders, 34,* 163-182.
Yaruss, J.S., & Quesal, R.W. (2004). Stuttering and the international classification of functioning, disability and health (ICF): an update. *Journal of Communication Disorders, 37,* 35-52.

APPENDIX A

"Therapy card" (translation of Dutch version)

Name:	Age:	Domicile:

Please fill in the percentage time spent on each treatment aspect.

DATE OF THERAPY SESSION → THERAPY PART↓					
Speech Motor Control (% time of the session)					
Emotion/cognition (% time of the session)					
Total %	100	100	100	100	100

Below, please mark those aspects of the therapy that were included in the specific therapy session. When treatment aspects cannot be classified into one of the given aspects, it should be classified as "others…"

SPEECH MOTOR CONTROL

GOAL	MEANS↓ DATE →					
(Continuing) Speak more fluent	• Articulation					
	• Continues Phonation					
	• Speech rate					
	• Intonation					
	• Loudness					
	• Others......viz.					
A. Physics, Coordination	Body coordination (independent from speech)					
	• Breathing					
	• Relaxation					
	• Phonation					
	• Others.......viz.					

COGNITION AND EMOTION

GOAL	MEANS ↓ DATE →					
A. Consioussness-raising and improvement of the insights of the stutter behaviour	Identification					
B. Decreasing the sensitiveness for stuttering behaviors and the reactions of others to it	Desensitization					
C. Decreasing negative cognitions and emotions and increasing the self-confidence and self-concept	See appendix					
D. Increasing the social skills	Assertiveness					
	Communication training/ Presentation training					
	Interaction training					
E. Others	....					

Chapter Four

Cross-cultural outcome evaluation of the ISTAR comprehensive stuttering program across Dutch and Canadian adults who stutter

Langevin, M., Huinck, W.J., Kully, D., Peters, H.F.M., Lomheim, H., & Tellers, M. (2006). Accepted for publication in the Journal of Fluency Disorders.

Abstract

There is a need to evaluate the effectiveness of stuttering treatment programs delivered in domestic and international contexts and to determine if treatment delivered internationally is culturally sensitive. Evaluation of the effectiveness of the ISTAR Comprehensive Stuttering Program (CSP) within and across client groups from the Netherlands and Canada revealed generally positive results. At 2 years post-treatment both groups were maintaining statistically significant reductions in stuttering frequency and improvements in attitudes, confidence, and perceptions as measured by the Revised Communication Attitude Inventory (S24), Perceptions of Stuttering Inventory (PSI), and the approach scale of the Self-Efficacy Scaling by Adult Stutterers (SESAS). Data pooled across the groups on these measures gave evidence of a global treatment effect with standardized effect sizes ranging from typical to larger than typical in the behavioural sciences. Only two differences between the groups emerged: differences in speech rate and perception of self. Given that these groups represent two distinct cultures, differences were discussed in terms of whether they could be due to cultural, methodological, or other variables. Overall, results suggest that, the CSP appears to be similarly effective in both cultures and thus, sufficiently sensitive to the culture of Dutch adults who stutter.

Introduction

In the face of increasing globalization and international sharing of information about stuttering and its treatment (see Pickering & McAllister, 2000; Shapiro et al., 2004), it is becoming ever more important for developers of treatment programs to demonstrate effectiveness, both in domestic and in international practice. In addition, international practice must be culturally sensitive. That is, drawing from Taylor (1986), Pickering and McAllister (2000), and Scollon and Scollon (2001), a treatment program for stuttering must be sensitive to the values and attitudes of those who stutter in the culture or subculture in which the treatment is being delivered and it must establish appropriate therapy goals, use appropriate materials, consider local therapy practices, and consider cultural influences on patterns of communication (*e.g.*, discourse systems, in particular functional uses of language; Scollon & Scollon).

In the context of this study culture refers to the "socially transmitted and shared influences" (Draguns, 1997, p. 214) on stuttering adults who are separated by geographical and language barriers (see Draguns). In terms of culture and communication disorders, Battle (1993) states that "Culture is about the behavior, beliefs, and values of a group of people who are brought together by their commonality" (p. xvii) and that "since the roots of communication are embedded in culture, it is logical to assume that one cannot study communication or communication disorders without reference to the cultural, historical, or societal basis of the communication style of the language used by members of the culture" (p. xix).[10]

[10] *Readers are referred to (a) Finn and Cordes (1997) for a discussion of the definition of culture; (b) Isaac (2002) for a discussion of cultural and linguistic diversity in the provision of speech-pathology services; and (c) Berry et al. (2002) for a discussion of the conceptions of culture.*

In response to the need to provide empirical evidence of the effectiveness of stuttering treatments in the Netherlands, an outcome evaluation project was undertaken by the second and fourth authors (Huinck & Peters, 2004). As part of that project, developers (the first and third authors) of the ISTAR Comprehensive Stuttering Program[11] (CSP; Boberg & Kully, 1985; Kully & Langevin, 1999; Langevin & Kully, 2003) were invited to participate in the Dutch outcomes project. This provided a unique opportunity for CSP developers to (a) evaluate the effectiveness of the CSP in the Netherlands, (b) compare the results with those obtained in Canada, and (c) consider whether any differences in outcome may be due to cultural differences between the two groups.

There is reason to believe that the CSP should be effective with linguistically and culturally diverse clients. Indeed, the CSP is an integrated treatment program that addresses speech characteristics and attitudinal issues that appear to be common to stuttering across cultures (see Finn & Cordes, 1997). As well, as Finn and Cordes note, similar behavioural treatments that teach "…speech fluency skills seem to be effective across a wide range of cultures and languages" (p. 229). Although the CSP has similarities to the Dutch adaptation (Franken, Boves, Peters, & Webster, 1992; Franken, Boves, & Peters 1997) of the Precision Fluency Shaping Program (PFSP; Webster, 1974), it also has differences. Whereas the PFSP focuses solely on the use of speech restructuring techniques (*e.g.*, fluency skills), the CSP integrates (a) speech restructuring and stuttering management techniques (*e.g.*, pullouts; Van Riper, 1973), which deal with core stuttering and learned struggle behaviours, and (b) cognitive-behavioural methods, which deal with the emotional and attitudinal aspects of stuttering.

The CSP is a treatment program for teens and adults that was developed at the Institute for Stuttering Treatment & Research (ISTAR) in Canada (Boberg & Kully, 1985). Although the general goals, clinical strategies, and components of treatment in the CSP are common to teens and adults, the *delivery* of treatment is age-and experientially-adjusted to meet the developmental levels of the two age groups. More recent descriptions of the program are provided in Kully and Langevin (1999) and Kully, Langevin, and Lomheim (in press). The CSP has been shown to be effective with adults and teens. In Boberg and Kully (1994), 76% of adults (13 of 17) and teens (19 of 25) respectively were maintaining satisfactory (< 3% percent syllables stuttered [%SS]) or marginally satisfactory (3.1–6.0 %SS) levels of fluency at 1 and 2 years post-treatment (adult means=1.33 and 0.96; teen means=1.11 and 1.26 %SS). In Langevin and Boberg (1993), 80% of adults (8 of 10) were maintaining satisfactory or marginally satisfactory levels of fluency at 1 year post-treatment (mean=1.3 %SS at 1 year follow-up). In addition, improvements in attitudes, confidence, and perceptions of speech performance were being maintained. That is, in Langevin and Boberg, at one year follow-up clients were maintaining an improvement of 7.2 raw scores on the Revised Communication Attitude Inventory (Andrews & Cutler, 1974), and improvements in percentage points of 23.5 on the Self Efficacy Scale for Adult Stutterers (Ornstein & Manning, 1985) and 33.1 on the Perceptions of Stuttering Inventory (Woolf, 1967). In both Langevin and Boberg and Boberg and Kully, 80% or more of clients rated their

[11] *The Comprehensive Stuttering Program is not related to the Swedish Comprehensive Stuttering Program (Forne-Wastlund, 2004) and is distinguished by specifying that it is the ISTAR program.*

speech as fair to good or generally good at follow-up and 87.5% or more felt like a normal speaker at least some of the time. The CSP has also been shown to be effective with clients who have a concomitant diagnosis of cluttering (Langevin & Boberg, 1996). However, the extent to which the CSP is effective in international practice is not known. If a treatment program yields similar results across cultures, the implication is that the program is culturally sensitive and that adaptations to treatment goals and processes are not required.

Finn and Cordes (1997) point out that it is not known if treatment variables will interact with cultural variables and, with the exception of a preliminary study that showed significant inter-judge differences in naturalness ratings of subjects with strong accents or dialects (Mackey, Finn, & Ingham, 1997), the role of cultural variables in evaluating stuttering treatment outcome is not known. For example, it was not known with certainty if the fluency skills used in the CSP would be appropriate for the Dutch language. It has been our clinical experience that fluency skill adjustments are needed for some non-English languages and their dialects to preserve naturalness. For example, in a Danish dialect that has frequent glottal stops, the "smooth blending" skill (similar to continuous airflow in other speech restructuring regimens) compromises naturalness. Similarly, Cooper and White [unpublished data, 1991, cited in Cooper & Cooper, 1993) indicate that "languages such as Spanish and German contain sound patterns that interfere with the use of some fluency eliciting vocal adjustments commonly used in America" (p. 202). Indeed, it was our experience that a minor adjustment of a fluency skill was needed to preserve naturalness within the Dutch language and across its dialects. Specifically, the 'g' in the southern area of The Netherlands is the voiceless velar phoneme [x], whereas in other areas it is the voiced uvular phoneme [χ]. As such, adjustments in the CSP "light touch skill" (soft contacts) had to be made for the differing dialects within the Dutch language to preserve naturalness. It is well known that natural sounding speech is a necessary outcome in order for clients to be prepared to use fluency skills outside of therapy. In the CSP naturalness is an ongoing goal in all speech practice throughout the program.

Further, there are also known differences in communication style between the Dutch and Canadian cultures in telephone practices. Whereas the Dutch answer the telephone or begin all telephone calls with their name, Canadians typically do not do so. This had implications for sequencing transfer practice, particularly given that many adults who stutter do so on their name. Sequencing transfer practice hierarchically is a fundamental clinical strategy in the CSP. Telephone transfer practice begins with telephone calls that are one sentence in length (*e.g.* a one question call to a business) and build systematically to conversation. Introducing oneself on the telephone in the treatment process in Canada does not occur until substantial mastery of skill use and confidence has been built to ensure success. Given the fear that telephone calls tend to stimulate, and the destabilization that can occur when success is not achieved, sequencing for success is particularly important. In the Netherlands, the introduction of self that was required for even one question telephone calls made the telephone transfer practice more difficult than is usually experienced in Canada.

Finally, it was not known if the tenets of the cognitive-behavioural component of the CSP would be completely appropriate, accepted, and practiced in The Netherlands. Again, its has been our clinical experience that that there are differences in the degree to which particular cognitive-behavioural skills are appropriate, accepted, and practiced by clients from non-English, non-Canadian Western and non-Western cultures. At present, however, we only have anecdotal evidence of such differences and the degree to which they may influence outcome in Western and non-Western cultures has yet to be investigated. Unfortunately, we do not as yet have measures to address these treatment process variables and their interaction with current outcome measures is not known. However, before launching an empirical investigation of the effect of clinically observed differences in treatment process variables across cultures, it is prudent first to evaluate outcome with traditional measures to determine if there are overall differences in outcome.

Given the paucity of cross-cultural investigations of treatment outcome in speech pathology, the broader literature of cross-cultural psychology was consulted. The study of cross-cultural psychology is concerned both with similarities and differences across cultures (Berry, Poortinga, Segall, & Dasen, 2002). Indeed, it is only against a background of similarity that cultural differences can be distinguished from alternative explanations of observed differences (see Campbell, 1964; Davidson, 1979). In the context of psychotherapy, Draguns (1997) suggests that the impact and pathways of cultural influence can be "invoked retrospectively" (p. 231). That is, differences between cultural groups in investigations are recorded and then traced back to plausible cultural characteristics that may have produced the differences. Thereafter, hypothesis-driven research across and within cultures is undertaken. Similarly, Van de Vijver and Lueng (1997) discuss two types of hypothesis testing studies in cross-cultural psychology.

The first kind of studies, *generalizability studies*, attempts to establish the generalizability of research findings obtained in one, typically Western, group to other Western or non-Western groups. In general, these studies make little or no reference to local cultural elements.

In the second type, called *theory-driven studies,* cultural factors are part of the theoretical framework. Cultural variation is deliberately sought as a validation of the model, and specific *a priori* predictions are proposed and tested. (p. 288).

As noted by Berry et al. (2002), interpretation of observed differences in generalizablity studies can only be made post hoc whereas, in theory-driven studies, alternative explanations for observed differences are ruled out.

In terms of cross-cultural methodology, this study is a generalizability study. There was no *a priori* selection of culture to which the Canadian developed CSP was to be compared and no *a priori* selection of a cultural variable upon which to make a hypothesis about its impact on treatment outcome. Indeed to do so would be premature given that we could not find in our literature a study that directly compares adult fluency treatment outcomes across cultures that are linguistically and geographically diverse. As stated earlier, it is very likely that no differences in outcome would be found given the universality of the speech and attitudinal aspects of stuttering and the previous history of successful speech restructuring treatment in The Netherlands. The reason for comparing the CSP in these two relatively similar cultures firstly was to empirically establish the effectiveness of the treatment in The Netherlands. That is, evidence of the effectiveness of the CSP must be available for critical appraisal in order for Dutch clinicians to determine if the CSP is a potential treatment for their clients, As well, in the Netherlands policy and financial decisions are based on evidence of the effectiveness of stuttering treatments.

The second compelling reason was to investigate whether or not the CSP appeared to be culturally sensitive. If differences in treatment outcome were found, the logical next step would be to determine whether or not such differences were related to cultural differences (*e.g.*, differences in language, values and attitudes, appropriate therapy goals, communication style, or response to treatment processes), or if such differences would be due to methodological or other variables. If it is found that treatment outcome is influenced by cultural differences, then the treatment program would need to be adapted.

Thus, given the importance of providing evidence for the effectiveness of stuttering treatments delivered domestically and internationally, in this case in The Netherlands, the primary purpose of this investigation was to evaluate the effectiveness of the CSP at 2 years post-treatment within and across client groups from the Netherlands and Canada and, if no differences between groups were found, to evaluate the global treatment effect across the groups. In view of the importance of delivering culturally sensitive treatment, a secondary purpose was to examine any observed differences and to consider whether they may be due to cultural, methodological, or other factors.

Method

Participants

Participants in this study were (a) 25 clients who were treated with the CSP in the Netherlands in 2000 and followed up 1 and 2 years post-treatment (Dutch group), and (b) 16 clients treated with the CSP who were part of a group of 18 clients treated in Canada in 1992 and followed each year for 5 years post-treatment (Canadian group). Two of the 18 Canadian clients were excluded from this investigation because they could not be contacted at 2 years follow-up; hence no 2 year follow-up data were available for these two cases. Four of the 16 Canadian participants were previously reported as individual case studies in Langevin and Kully (2004). In the Langevin and Kully study, the raw speech and questionnaire data for these 4 participants were used to illustrate individual differences and trends over the 5 year follow-up period. In contrast, the present study used the 2 year post-treatment data for the 4 participants as part of the group analyses to compare group means, ranges, effect sizes, and proportions of the Dutch group and the Canadian group. In essence, in the Langevin and Kully study, data from these 4 participants were used as exemplars of individual trends while in the present study, the data were used as part of an overall group analysis.

Dutch Group

Of the Dutch group (mean age=29.6 years; range=17-53; males=17; females=8) 56% (14) of the participants had post-secondary (*i.e.*, university) education, the remainder had high school education.[12] Dutch was the first language for all participants and

[12] *Dutch educational achievement levels were equated to Canadian levels because it was impossible to do the reverse.*

English was a second or third language for 23 participants. To be included in this study, participants had to meet the following criteria: reported onset of stuttering before 6 years of age; no reported problems in motor development, speech development other than stuttering, language development, or hearing; no reported use of medication that could influence respiration, phonation, or articulation; and no reported psychiatric problems. Receipt of therapy was contingent on participation in the Dutch outcomes research project (Huinck & Peters, 2004). Forty-four percent (11) of participants reported having previously taken one type of therapy, 28% (7) had previously taken two different types of therapy, and 28% (7) had previously taken three or more (range 3-6) different types of therapy programs. Therapy types were defined as therapy programs that were delivered in school during primary and secondary school years and therapy programs that were delivered in individual formats (*e.g.*, over a period of 1 year with a therapist in a hospital setting, private practice, or fluency specialist clinic) and in intensive formats (*e.g.*, 10 days to 24 days of therapy that were either consecutive or delivered in clusters). Within each type of therapy, the therapeutic approach targeted either speech modification through speech restructuring and/or stuttering modification, attitudes and emotions associated with stuttering and avoidance reduction, or combinations of both. For example, one Dutch client who had received four types of therapy had had school based therapy, individual therapy, and two different intensive programs. Through participation in these types of programs the participant had had experience with speech restructuring, stuttering modification, attitude-emotion change, and avoidance reduction. Intensive programs that were repeated were considered one type of therapy. All but one of the participants had not had therapy for a period of 5 years prior to the CSP.

Canadian Group

Of the Canadian group (mean age=24.6 years; range=15-42 years; males=13; females=3), 43.75% (7) of participants had post-secondary education; the remainder had high school education. All participants had reported onset of stuttering in childhood; none reported psychiatric problems or problems in motor or language development and none reported use of medications that could affect therapy. One participant had a concomitant diagnosis of cluttering and one participant reported having a severe congenital hearing loss in one ear. One participant had no previous therapy, 2 had limited therapy within the year prior to the CSP, 1 had therapy 1 year prior to the CSP, and the remainder had had therapy but not for 2 to 10 years prior to the CSP. Of the 15 participants who had previous therapy, 75% (12) reported having previously received one type of therapy and 25% (3) had received two different types of therapy. Only one participant had previously received the CSP intensive therapy program under investigation. Although enrolment in the CSP was not contingent upon participation in this outcome evaluation, all participants had consented to be contacted in the post-treatment period.

Clinical Program

The CSP evaluated in this study was a three week intensive group-therapy program for adults. Clients received 90 hours of therapy (6 hours per day). Therapy consisted of a combination of individual, small-group, and large-group activities that targeted speech restructuring, stuttering management, self-management goals (*e.g.*, self-monitoring, self-evaluation, and problem-solving skills), and attitudinal-emotional change. The clinician-client ratio varied over the course of the treatment day from 1:1 to 1:3 depending on the needs of the participants and the demands of the activity. Client-clinician and client-client pairings were rotated daily to facilitate generalization.

The CSP has three phases: acquisition of fluency and cognitive behavioural skills (weeks 1 and 2), transfer (week 3), and maintenance (continued practice of speech restructuring, stuttering management, and cognitive-behavioural skills in the months and years following the 3 week intensive program). During acquisition, clients learn fluency enhancing techniques (*e.g.*, easy onsets, soft contacts) that are taught within a framework of prolonged speech, stuttering management skills (*e.g.*, tension modification and pullouts) to manage oral and laryngeal tension and residual stuttering, and self-management skills (*i.e.*, self-monitoring, self-evaluation, and problem solving). Clients also are introduced to cognitive-behavioural skills to deal with attitudes (*i.e.*, affective, cognitive, and behavioural responses) associated with stuttering. In addition to speech management goals (*i.e.*, improved fluency and improved management of tension and residual stuttering) attitudinal goals are implemented to help clients (a) achieve improved communication, social skills, and confidence, (b) develop positive attitudes toward communication, (c) reduce avoidances, (d) develop acceptance of and openness about stuttering and about fluency skills (e) develop the ability to manage fear and anxiety and deal with negative listener reactions, and (f) manage regression and recognize when relapse is occurring. During the transfer phase, clients use speech and cognitive behavioural skills in a hierarchy of transfer activities that moves from least to most difficult and includes standard tasks (*e.g.*, talking on the telephone, talking to groups) and personalized tasks (*e.g.*, teaching a class, giving a business presentation). Preparation for maintenance begins in acquisition but is more intensive during the transfer phase. In preparation for maintenance, clients are encouraged to (a) carry out home practice of fluency skills and continue transfer activities, (b) make lifestyle adjustments to expand speaking opportunities, (c) join or form a self-help group, and (c) seek follow-up therapy as needed. Self management strategies are integral to all phases of the treatment program and training continues throughout the program. More detailed descriptions of the CSP are available in Kully and Langevin (1999) and Kully, Langevin, and Lomheim (in press).

At ISTAR the CSP is delivered by a clinical team that includes (a) senior ISTAR staff, (b) student speech-language pathologists, and (c) practicing speech-language pathologists who wish to obtain specialized experience in treating adults who stutter. Prior to the intensive clinic, both student and practicing speech-language pathologists attend a CSP training workshop. To ensure treatment integrity for the Dutch group, the CSP was delivered by a team that included senior ISTAR staff (the first and fifth authors) and a Dutch clinical coordinator (the sixth author) who had previously been trained at ISTAR. The team also included eight Dutch trainees, half of whom had had previous experience with the CSP at ISTAR. With the exception of the introductory cognitive-behavioural seminars which were delivered in English, all other treatment was delivered in Dutch. For clients who needed translation of the introductory seminars, simultaneous translation was carried out by the Dutch clinical coordinator (who is proficient in both languages and has extensive experience in stuttering treatment). Client handouts were translated into Dutch by professional translators and reviewed by the Dutch Clinical coordinator. In some instances practice text (*e.g.*, sentences used for fluency skill practice) was translated by the Dutch clinical coordinator. In contrast to therapy delivered at ISTAR, consisting of 8 to 12 clients per intensive clinic, the Dutch program was essentially a double clinic (*i.e.*, two Dutch groups were treated simultaneously). However, the clinician-client ratio was the same as that at ISTAR. Also, due to difficulties obtaining accommodation for Dutch participants in the City of Nijmegen, the Dutch program was delivered as a residential program (*i.e.*, the facility housed all participants) in a small rural community. However, during the transfer phase, clients were transported to the City of Nijmegen to carry out transfer tasks. At ISTAR, which is located on the University Campus in the centre of the City of Edmonton, clients are housed in various on-campus locations of their own choosing. Despite differences in housing arrangements, there was no difference between the groups in the number of hours spent in therapy. With the exception of the minor adjustments needed in the light touch skill in the Dutch group, all other aspects of the treatment methodology were the same across the Dutch and Canadian groups.

Treatment Outcome Measures

Speech Measures

The dependent speech measures were percent syllables stuttered (%SS) and syllables spoken per minute (SPM). Speech samples were obtained pre-treatment (Pre), immediately post-treatment (Post), and at one- (F1) and two-years (F2) follow-up. For the Dutch group, within- and beyond-clinic measures were obtained at all measurement times. For the Canadian comparison group only beyond-clinic measures were obtained. Historically, within-clinic follow-up measures have not been possible to obtain at ISTAR because the majority of clients live hundreds to thousands of miles away from the clinic. For the Dutch within-clinic measures, each participant was video-recorded for 3 minutes of talking time in an interview, a reading, and a monologue task. Accumulation of talking time for the interview began 30 seconds after the interview commenced. All participants were asked the same set of questions which were related to stuttering. With respect to the monologue, participants chose a topic from a set of topic cards (*e.g.*, vacation, hobbies etc.). Topics sets differed at each measurement session. Recordings were made in a building not associated with treatment by an unfamiliar research assistant who was not associated with the therapy program. For the Dutch and Canadian beyond-clinic measure, each participant was audio-recorded for 2 minutes of talk time while speaking on the telephone. Pre- and post-treatment samples were of telephone calls made to businesses. Follow-up samples were of surprise phone calls received from unfamiliar research assistants. After a standard introduction by the research assistant, participants and research assistants conversed about topics that were relevant to the participant. Surprise telephone samples were chosen as follow-up measures because they are generally considered to be among the most difficult speech situation for many people who stutter (Boberg & Sawyer, 1977; Bloodstein, 1987; O'Brian, Onslow, Cream, & Packman, 2003), they represent real life speech performance, and they are relatively free of clinic cues.

Reliability of speech measures

Speech measures were obtained from analysis of the video- and audio-recorded speech samples. Frequency counts of stutters and syllables spoken were made on an electronic button-press event recorder with a timing device. Frequency counts using this methodology have shown it to be a reliable measure of stuttering (*e.g.*, Boberg & Kully, 1985; Langevin & Boberg, 1993; Lincoln, Onslow, Lewis, & Wilson, 1996; O'Brian et al., 2003). Counting guidelines established at ISTAR (Kully, 1986) were used. The SPM measures were of overall speech rate rather than articulatory rate (Ingham & Riley, 1998). Extraneous utterances such as interjections were excluded and each syllable was counted once only, regardless of the number of times it was repeated. Lengthy meditative pauses (*i.e.*, greater than 2 seconds) were excluded. Three Dutch and four Canadian research assistants who were independent of the therapy program's administration were trained using a program that involved six hours of training in counting stuttered and fluent syllables from video-taped and audio-taped speakers who stutter. In addition to initial instruction with the traditional English rater training program (Boberg & Kully, 1994; Langevin & Boberg, 1993), Dutch research assistants were trained using samples of Dutch speakers who stutter. Both sets of raters were trained to criterion before making ratings. Speech samples in this study were randomly assigned to raters; however, in the Dutch group samples were stratified for participants (*i.e.*, one rater rated a participant's samples from all 4 measurement occasions, but did so without identifying information).

To obtain inter-rater reliability estimates, fifty-two Dutch speech samples (13 from each of Pre, Post, F1, and F2 for 13 randomly chosen participants), representing 11% of the Dutch participants' data set, were rated by all three Dutch raters. Similarly, eight Canadian speech samples (2 from each of Pre, Post, F1, and F2 for 2 randomly chosen participants), representing 14.8% of the Canadian data set in this study, were selected for inter-rater reliability. Because one Canadian rater was not available to rate samples at F2, reliability measures were calculated for 3 raters on 8 samples (*i.e.*, samples included F2 measures) and 4 raters on 6 samples (*i.e.*, samples did not include F2 measures). The intra-class correlation coefficients[13] (two-way mixed effects model) for inter-rater reliability for Dutch and Canadian raters were equal to or exceeded 0.97 for %SS and 0.93 for SPM.

Speech Naturalness Ratings

To give context for the interpretation of F2 speech outcome measures for each group, naturalness (NAT) ratings are included in this report. However, comparison of NAT ratings cannot be made because different scales and procedures were used in each country. In both countries the participants in this study are also participants in larger naturalness rating studies that are underway. Each of the studies has a different purpose, hence different methodologies were used.

For the Dutch group the seven-point bipolar naturalness scale from the perceptual rating instrument developed by Franken, Boves, Peters, and Webster (1992; 1995) and Franken, Boves, and Peters (1997) was used. The bipolar naturalness scale is defined by contrastive terms labelling extremes (*e.g.*, *1=very unnatural; 7=very natural*). Naturalness ratings were obtained for 13 randomly chosen participants representing 52% of the Dutch group. These restrictions were prompted by practical reasons (containment of the duration of the rating session to prevent the raters from becoming overtaxed). Forty-five second speech samples were extracted from the F2 within-clinic monologue speech samples.

---

[13] *Intraclass correlation coefficients range from 0.00 to 1.00.*

Selection of samples was not based on presence or absence of stuttering, that is, samples may or may not have contained stutters. These samples, along with 143 others that were related to the larger naturalness study, were presented to raters in one session. Ratings were made by forty-two unsophisticated female listeners (mean=20.2 years of age; range=17-26 years of age) who were first year logopaedics students. The rating session took 3 hours, inclusive of instruction time and three breaks. Reliability of the scale was 0.95 using an intra-class correlation coefficient (two-way mixed effects model).

For the Canadian group, the nine-point rating scale developed by Martin, Haroldson, and Triden (1984) was used (*1=highly natural, 9=highly unnatural*). Naturalness ratings of F2 speech samples were obtained for 8 participants in the Canadian group representing 57% of the participants in this study for whom F2 speech samples were obtained. These 8 participants had been selected for the larger Canadian longitudinal study on the basis of longitudinal data availability (*i.e.*, 4 to 5 years of speech data were available for use in the larger naturalness study). Fifteen-second stutter-free speech samples were drawn from F2 beyond-clinic telephone samples. To ensure samples were perceptually free of stuttering, all samples were independently analysed by the first and third author and a research assistant at ISTAR who was trained in identifying and analysing stuttered and fluent speech. No samples were judged to contain instances of stuttering. Ratings were made by 30 unsophisticated listeners (mean age=27.33 years; range=18–52; females=28; males=2). Raters were seven first-year and fourteen second-year graduate speech language pathology students and nine community members. Speech samples were presented on one of two tapes containing these and 140 other speech samples. Rating sessions were conducted individually or in small groups of up to 5 raters. The rating sessions took approximately 2 hours inclusive of instructions, ratings, two 10-minute breaks, and debriefing. Reliability of the scale was .92 using an intra-class correlation coefficient (two-way mixed effects model).

Mean naturalness ratings for each group were derived from the means of listener ratings for each participant in each group and were compared to the normative data previously reported for each scale. That is, Franken et al. (1992) reported a mean of 5.06 (SD=0.50) for Dutch non-stutterers. For the Martin et al. (1984) scale, reported means for non-stutterers have ranged from 2.3 (SD=1.21; Martin & Haroldson, 1992) to 3.6 (SD=2.1; O'Brian et al., 2003).

Clinically Meaningful Maintenance

To determine the extent to which Dutch and Canadian participants did and did not maintain clinically meaningful gains and to compare results, a methodology was devised that considered stuttering frequency in follow-up that is relative to both Pre *and* Post measures. That is, participants were categorized as *not* maintaining clinically meaningful speech gains (Non-Maintainers) if they were not maintaining at least 50% improvement in %SS at F2 relative to their Pre %SS *and* if they were not maintaining an F2 %SS that was equal to or less than their Post %SS plus 3% (*i.e.*, participants were allowed 3% regression in follow-up relative to post-treatment).

This methodology is a preliminary attempt to characterize maintenance of speech gains.

It takes into account the inherent variability of stuttering, the problem in defining successful treatment outcome for those with low pre-treatment %SS scores (and therefore, truncated ranges of improvement in follow-up relative to pre-treatment), and the improbability that treatment will completely remove stuttering in adolescents and adults. It is also used in the absence of much needed research to develop a model of clinically meaningful maintenance or clinically meaningful treatment change that considers input from clients, clinicians, researchers, and significant others (Finn, 2003). It is a modest step beyond that which was used in previous CSP outcome evaluations (Boberg & Kully, 1994; Langevin & Boberg, 1993) and definitions that consider 2-3 %SS in follow-up as being reflective of relapse (*e.g.*, Andrews & Craig, 1988; Blood 1995; Craig et al., 1996).

In this investigation the methodology used to determine clinically meaningful maintenance was applied to the beyond-clinic measure obtained at F2 for the Dutch and Canadian participants (*i.e.*, the surprise telephone calls). Thereafter the methodology was applied to the Dutch within-clinic measures (*i.e.*, interview, reading, and monologue) at F2 to determine the degree to which there was agreement across the Dutch within- and beyond-clinic conditions.

Self Report Measures

Self-report measures were (a) the Revised Communication Attitude Inventory (S24; Andrews & Cutler, 1974); (b) Perceptions of Stuttering Inventory (PSI; Woolf, 1967), which measures clients' perceptions of struggle and avoidance behaviours associated with stuttering and the expectancy to stutter; (c) the approach scale of the Self-Efficacy Scaling by Adult Stutterers (SESAS; Ornstein & Manning, 1985), which measures clients' confidence in entering a variety of speaking situations, and (d) the Speech Performance Questionnaire, which measures clients' perceptions of post-treatment speech performance (SPQ; adapted from Perkins, 1981).[14]

Data Analysis

PSY software (Bird, Hadzi-Pavlovic, & Isaac, 2000) was used to calculate point and interval estimates of raw effect sizes (scaled in dependent variable units; *e.g.*, %SS) and standardized effect sizes (Cohen's *d*; 1988). The interval estimates are simultaneous confidence intervals (CIs) for all Pre-F2 contrasts.[15]

---

[14] *Challenges to the validity of the S24 by Ulliana and Ingham (1984) and the concerns raised by Hillis and McHugh (1998) regarding the interdependence of efficacy expectation and performance accomplishments have been discussed in Langevin and Kully (2003).*

[15] *Although Algina and Keselman (2003) have shown that Bird's (2002) approximate confidence intervals for standardized effects sizes tend to be liberal (i.e. they may have coverage probabilities less than 0.95), their suggestions proposed for better approximate CIs deal only with one-way between or one-way within designs and could not accommodate the two-way between and within design needed to analyse the pooled data in this study. When suggestions were applied for the one-way within analyses, differences in the magnitude of the confidence intervals did not change the interpretation (e.g., an upper limit that reflected a large effect did so with and without the Algina and Keselman suggestion).*

Following suggestions of Betz and Levin (1982) and Bird and Hadzi-Pavlovic 2005), the per-contrast Type 1 error rate for these measures was set at 0.025.[16] Calculation of the raw effect size is the same as that for Cohen's *d* (*i.e.*, $\{$mean $1_{(Pre)}$ – mean $2_{(F2)}\}/SD_{pooled}$). Estimates of raw effect sizes are reported in addition to Cohen's *d* because they are readily interpretable by clinicians and researchers (see Bond, Wiitala, & Richard, 2003; Kraemer et al., 2003). For example, a raw effect size of 6.86 in a Pre-F2 %SS contrast indicates that on average participants were maintaining a reduction of 6.86 %SS at F2. Confidence intervals provide information about the magnitude of a contrast and the precision of estimation to the population (Bird, 2002). A broad interval around an effect size indicates a less precise estimate of what the effect size would be in the population. A statistically significant pooled Pre-F2 contrast reflects the global treatment effect across the groups.

The use of standardized effect sizes facilitates the interpretation of results. However, Kraemer et al. (2003) advise that interpretation should be made relative to typical effects found in the relevant research literature. In the absence of established benchmarks for *d* effect sizes in long-term stuttering treatment outcome, the suggestions of Kraemer et al. were followed. Kramer et al. state that Cohen's benchmarks of small (0.20), medium (0.50) and Large (0.80) were meant to be relative to typical findings in *behaviour research in general* and as such give guidelines for interpreting *d* effects that are relative to effect sizes typically found in applied behavioural research. That is, an effect size of 0.20 is small or smaller than typical in the applied behavioural sciences, 0.50 is medium or typical, 0.80 is large or larger than typical, and ≥ 1.0 is much larger than typical. They also suggest that a range of values be used to interpret *d*. For example, a *d* greater than 0.90 would be described as much larger than typical in the applied behavioural sciences, a *d* of between 0.70 and 0.90 would be described as larger than typical, and a *d* of between 0.60 and 0.70 would be described as typical to larger than typical.

Results

Follow-up Contact and Questionnaire Response Rates

Dutch group

Remarkably, within-clinic speech measures were obtained from 100% of the Dutch participants at F1 and F2. Beyond-clinic measures were obtained from 76% of participants (19 of 25) at F1 and 96% (24 of 25) at F2. The S24, SESAS, and PSI were

---

[16] *The analysis of data from the beyond-clinic measures was based on a simple effects model rather than the two-factor ANOVA model in order to allow for the inclusion of simple as well as main and interaction effect contrasts involving pre-post differences. Betz and Levin (1982) and Bird and Hadzi-Pavlovic (2005) recommend that the expected number of Type I errors from such an analysis should be set at 2α, the expected overall error rate from a conventional two-factor ANOVA-model analysis including only main and interaction effect contrasts involving pre-post differences. The Bonferroni-adjusted per-contrast error rate was set at 0.025 to ensure that the overall error rate for inferences on each dependent variable cannot exceed $2\alpha=0.10$ if the nominal error rate is $\alpha=0.05$. The same per-contrast error rate (0.025) was used for analyses of data from the within-clinic measures, although the rationale for this choice was to maintain the same (slightly conservative) per-contrast error rate for all contrasts on all measures.*

obtained from 100% of participants at F1 and 92% (23 of 25) at F2. Speech performance questionnaires were returned by 92% (23 of 25) at each of F1 and F2.

Canadian group

Of the 16 participants in the Canadian group, 2 year post-treatment speech data were available for 14 (which represented 78% of the total group of 18) and questionnaire data were available for 10 (which represented 56% of the total group of 18). Two of the participants who returned questionnaires at 2 years post-treatment could not be contacted by telephone to obtain speech measures.

Table 1. Descriptive statistics for percent syllables stuttered (%SS) and syllables per minute (SPM) at pre-treatment (Pre), post-treatment (Post), and 1- and 2-years follow-up (F1 and F2) for the within clinic measures for the Dutch group and the beyond clinic measure for the Dutch and Canadian groups.

Measure/ Occasion		*%SS*			*SPM*		
	N	*M*	*SD*	*Range*	*M*	*SD*	*Range*
Within-clinic							
Interview							
Pre	24	10.43	10.63	0.39-43.56	132.09	51.61	46.32-210.88
Post	25	1.16	2.34	0-9.20	169.66	31.72	112.77-224.72
F1	25	7.57	10.45	0.23-47.78	142.01	41.22	53.14-214.21
F2	25	6.78	9.90	0.19-39.17	150.25	48.80	39.80-214.29
Reading							
Pre	25	13.56	15.10	0-59.23	135.75	60.30	32.41-269.59
Post	25	0.75	2.02	0-9.69	162.88	35.30	94.87-231.27
F1	25	5.57	11.85	0-50.00	161.53	55.92	44.16-251.57
F2	25	3.60	6.87	0-28.74	169.27	56.37	36.50-238.50
Monologue							
Pre	25	12.56	11.62	0.33-45.74	121.19	50.47	30.50-207.21
Post	25	1.18	2.46	0-8.88	163.34	39.54	94.59-245.45
F1	25	7.57	13.89	0-67.33	146.24	44.34	48.57-229.90
F2	25	6.25	9.30	0-41.22	148.67	47.47	43.96-232.17
Beyond-clinic							
Dutch							
Pre	25	12.00	10.73	1.52-42.76	145.33	51.98	62.61-256.73
Post	25	3.24	5.25	0-17.46	175.84	38.72	111.64-247.83
F1	19	6.63	10.74	0-44.30	148.89	53.23	37.83-224.32
F2	24	7.04	8.99	0.22-36.52	134.51	48.36	43.79-232.47
Canadian							
Pre	14	11.99	5.72	3.70-22.10	130.49	35.81	46.00-191.30
Post	14	0.91	0.83	0-2.80	146.38	18.27	112.30-182.60
F1	12	3.88	5.70	0-20.40	153.44	27.75	112.70-195.10
F2	14	4.38	7.31	0-29.20	153.36	34.32	49.40-186.60

Speech Measures
Table 1 presents the descriptive statistics for %SS and SPM for the within-clinic measures for the Dutch group and the beyond-clinic measures for the Dutch and Canadian groups.

Stuttering frequency
As shown in Table 1, the Dutch group achieved substantial reductions in stuttering frequency in the within-clinic measures at post-treatment with means that ranged from 0.75 to 1.18. These reductions are comparable to those that are typically obtained in immediate post-treatment within-clinic measures at ISTAR (*e.g.*, 0.53 %SS in a video-taped conversation in Langevin & Boberg, 1993). However, the mean of 3.24 %SS for the Dutch group in the beyond-clinic Post measure was at minimum 1.95 %SS higher than has been reported for Canadian groups in this study and others. That is, across this study, Langevin and Boberg (1993), and Boberg and Kully (1994), immediate post-treatment means for adults ranged from 0.80 to 1.29. Inspection of individual Post data revealed that three Dutch participants accounted for the elevated levels of stuttering frequency at post-treatment. These three participants had beyond-clinic scores that ranged from 15.05 to 17.46 %SS. This unusual post-treatment response also occurred in their within-clinic measures. In contrast, for 19 Dutch participants the Post range was 0-2.8 %SS. For the remaining 3 Dutch participants, the Post range was 2.9 to 6.31 %SS. As shown in Table 1, the Post range for the Canadian group was 0-2.8%.

Table 2. Raw unit and standardized effect sizes at two years follow-up for percent syllables stuttered (%SS) and syllables per minute (SPM) for the Dutch within- and beyond-clinic measures, the Canadian beyond-clinic measure, and pooled beyond-clinic data.

Group (n)	*t*	*Raw Effect Size*				*Standardized Effect Size*			
		Effect Size	*SE*	*CI Lower*	*CI Upper*	*d*	*SE*	*CI Lower*	*CI Upper*
%SS									
Within-clinic[a,b]									
Dutch (25)	5.46 **	6.86	1.26	3.85	9.86	0.67	0.12	0.38	0.96
Beyond-clinic[c]									
Dutch (24)	3.59**	4.61	1.29	1.60	7.62	0.52	0.15	0.18	0.86
Canadian (14)	4.52**	7.62	1.68	3.68	11.55	0.86	0.19	0.42	1.30
Canadian-Dutch	-1.42	-3.01	2.12	-7.96	1.95	-0.34	0.24	-0.90	0.22
Pooled (38)	5.77**	6.11	1.06	3.64	8.59	0.69	0.12	0.41	0.97
SPM									
Within-clinic[b]									
Dutch (25)	-4.37**	-26.74	6.12	-41.38	-12.11	-0.58	0.13	-0.89	-0.26
Beyond-clinic[c]									
Dutch (24)	1.94	12.37	6.37	-2.55	27.40	0.27	0.14	-0.06	0.60
Canadian (14)	-2.74*	-22.87	8.35	-42.40	-3.35	-0.50	0.18	-0.93	-0.07
Canadian-Dutch	3.35*	35.24	10.50	10.67	59.82	0.77	0.23	0.23	1.31
Pooled (38)	-1.00	-5.25	5.25	-17.54	7.04	-0.12	0.12	-0.39	0.16

Note. [a]*Data were positively skewed and were therefore also analysed using a log transformation: d = 0.87 (SE=0.14; CI = 0.54,1.20);* [b]*df=24;* [c]*df=36; * p< 0.01, **p<0.001.*

Raw and standardized effect sizes for %SS at F2 (*i.e.*, the Pre-F2 contrast) are presented in the upper portion of Table 2. The Dutch within-clinic effect size was calculated using the means of the within-clinic measures.

All %SS contrasts were statistically significant for the Dutch within- and beyond-clinic measures and the Canadian beyond-clinic measure. It is notable that the CIs for the Dutch beyond-clinic measure are more precise (*i.e.*, more narrow) than that for the Canadian group. However, this is in part due to the larger samples size for the Dutch group. As would be indicated by the degree of overlap between the CIs for these two data sets, there was no significant difference between the groups in the magnitude of the Pre-F2 contrast (see Canadian-Dutch, Table 2) and the pooled contrast was statistically significant. The narrower CI for the pooled data reflects greater accuracy of measurement, and the small standard error indicates a relatively precise estimate of the magnitude of the treatment effect across the groups.

Clinically meaningful maintenance

Recall that participants were categorized as Maintainers and Non-Maintainers according to their F2 %SS data (*i.e.* n=24 for the Dutch group and n=14 for the Canadian group). As a result there were 71% (17) Dutch Maintainers (F2 mean %SS=3.32, SD=3.78) and 29% (7) Non-Maintainers (F2 mean %SS=16.07, SD=11.72). In comparison there were 86% (12) Canadian Maintainers (F2 mean %SS=2.38; SD=1.64) and 14% (2) Non-Maintainers (F2 mean %SS=16.35; SD=18.17) respectively. To determine if there would be agreement across the within- and beyond-clinic measurements for the Dutch group, the methodology to categorize Maintainers and Non-Maintainers was applied to the F2 Dutch within-clinic measures. Results indicated that performance in the within-clinic measures mirrored that of the beyond-clinic measure at F2 for 94% (16 of 17) of the Maintainers and 71% (5 of the 7) Non-Maintainers.[17]

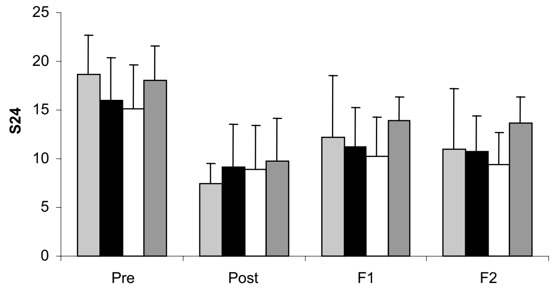
Speech rate

As shown in Table 1, speech rates for the Dutch group increased in all within-clinic measures; however, in the beyond-clinic measure the increase in speech rate at Post (mean=175.84) was not maintained at F2 (mean=134.51). Surprisingly, the F2 speech rate was lower than at Pre (mean=145.33). For the within-clinic measure, treatment effects were calculated using the mean of the three within-clinic speech rate measures. As shown in the lower portion of Table 2, the within-clinic Pre-F2 contrast was statistically significant. In the beyond-clinic measure the Pre-F2 contrast was not significant for the Dutch group, but was significant for the Canadian group. Accordingly, the Dutch-Canadian Pre-F2 contrast was statistically significant, indicating that there was a difference between the groups, and the pooled data contrast was non-significant.

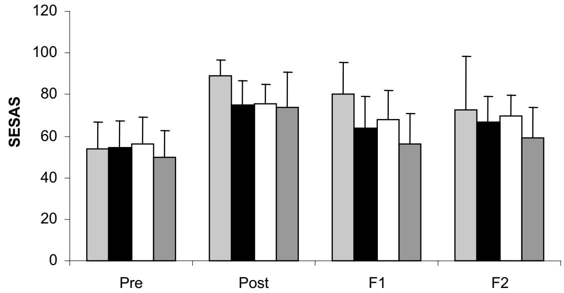
[17] *Nine Dutch Maintainers (53%) also were maintaining on all 3 within-clinic measures and 7 (41%) were maintaining on 2 of the 3 within-clinic measures. One Maintainer was maintaining only on the beyond-clinic condition. For the Non-maintainers, 4 (57%) were not maintaining on all within-clinic measures, 1 (14%) was not maintaining on 2 within-clinic measures, and 2 were not maintaining on only the beyond-clinic measure.*

Naturalness Ratings
Recall that *higher* scores for the Dutch group (the seven-point, unnatural-natural bipolar scale) and *lower* scores for the Canadian group (the nine-point natural-unnatural scale) reflect speech that was perceived as more natural. For the Dutch group the mean NAT rating at F2 was 4.03 (SD=0.79; Median=4.17; Range=2.69-5.19). These data indicate that the Dutch NAT ratings were 1.03 scale scores lower than the mean for non-stutterers (5.06) reported by Franken et al. (1992). For the Canadian group, the mean NAT rating at F2 was 2.85 (SD=0.73; Median=2.86; Range=1.70-3.77). This NAT rating was within the range of mean NAT ratings previously reported for non-stutterers; that is from a low of 2.3 in Martin & Haroldson (1992) to a high of 3.6 in O'Brian et al., (2003).

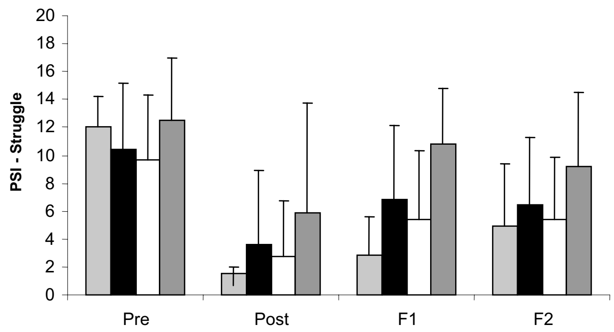
Self-Report Measures
S24, SESAS, and PSI subscales. Figures 1A-E present the means and standard deviations for the S24, SESAS, and the subscales of the PSI for the Canadian and Dutch groups. Table 3 presents the raw and standardized effect sizes for all Pre-F2 contrasts. All contrasts for the S24, SESAS, and PSI subscales were statistically significant for each of the Dutch and Canadian groups. Although the magnitudes of the standardized effect sizes were generally larger for the Canadian group, in part due to smaller sample size, the CIs were more precise for the Dutch group. As indicated by the degree of overlap between the CIs for the Dutch and Canadian groups across the measures, there were no significant differences between the groups in the Pre-F2 contrasts (see Canadian-Dutch contrasts) and the pooled Pre-F2 contrasts were statistically significant.



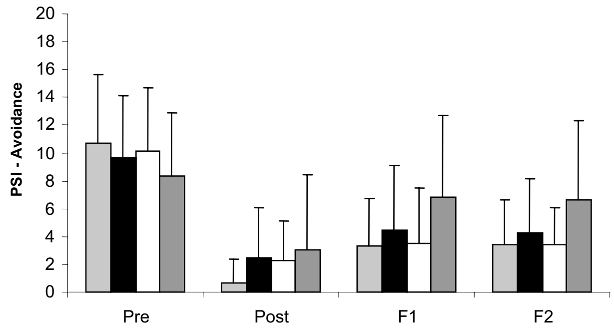
(a)



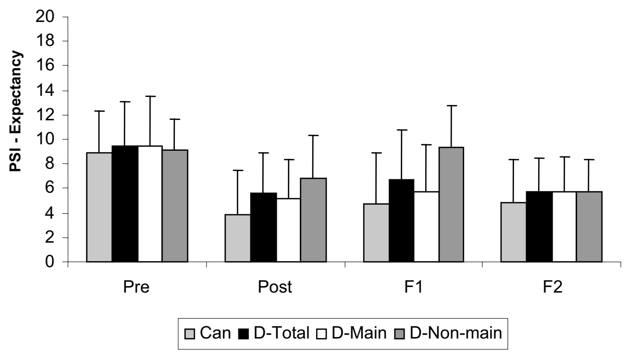
(b)



(c)



(d)



(e)

Figure 1. Means and standard deviations (error bars = + 1SD)[a] for the Revised Communication Attitude Inventory (S24), Self-Efficacy Scaling for Adult Stutterers (SESAS), and the Struggle, Avoidance, and Expectancy subscales of the Perceptions of Stuttering Inventory (PSI) for the Canadian group (Can), the Dutch group as a whole (D-Total), and for the Dutch Maintainers (D-Main) and Non-Maintainers (D-Non-Main). Note: For clarity -1SD errors bars are not shown.

Table 3. Raw usnit and standardized effects sizes at two-years follow-up for the Revised Communication Attitude Inventory (S24), the approach scale of the Self-Efficacy Scaling by Adult Stutterers (SESAS), and the Struggle, Avoidance, and Expectancy subscales of the Perceptions of Stuttering Inventory (PSI) for the Dutch and Canadian groups.

		Raw Unit Effect Sizes[a]				*Standardized Effect Sizes*			
Measure/Group (n)	*t*	*Effect Size*	*SE*	*CI Lower*	*CI Upper*	*d*	*SE*	*CI Lower*	*CI Upper*
S24[b]									
Dutch (23)	5.13**	5.04	0.98	2.73	7.36	1.15	0.22	0.62	1.67
Canadian (10)	5.30**	7.90	1.49	4.39	11.41	1.80	0.34	1.00	2.59
Canadian-Dutch	-1.60	-2.86	1.78	-7.06	1.35	-0.65	0.41	-1.60	0.31
Pooled (33)	7.25**	--6.47	0.89	4.37	8.57	1.47	0.20	0.99	1.95
SESAS[b]									
Dutch (23)	-4.09**	-13.42	3.28	-21.13	-5.70	-0.91	0.22	-1.43	-0.39
Canadian (10)	-3.73**	-18.53	4.97	-30.24	-6.83	-1.25	0.34	-2.05	-0.46
Canadian-Dutch	0.86	5.11	5.95	-8.91	19.13	0.35	0.40	-0.60	1.30
Pooled (33)	-5.37**	-15.97	2.98	-22.98	-8.96	-1.08	0.20	-1.56	-0.61
PSI (Struggle)[c]									
Dutch (23)	3.77**	4.00	1.06	1.49	6.51	0.90	0.24	0.34	1.46
Canadian (10)	4.29**	7.11	1.66	3.19	11.03	1.60	0.37	0.72	2.48
Canadian-Dutch	-1.58	-3.11	1.97	-7.76	1.54	-0.70	0.44	-1.74	0.35
Pooled (33)	5.65**	5.56	0.98	3.23	7.88	1.25	0.22	0.73	1.77
PSI (Avoidance)[c]									
Dutch (23)	3.73**	5.41	1.14	2.71	8.11	1.30	0.27	0.65	1.94
Canadian (10)	4.04**	7.22	1.79	2.99	11.45	1.73	0.43	0.72	2.74
Canadian-Dutch	-0.85	-1.81	2.12	-6.83	3.21	-0.43	0.51	-1.64	0.77
Pooled (33)	5.95**	6.32	1.06	3.81	8.83	1.51	0.25	0.91	2.11
PSI (Expectancy)[c]									
Dutch (23)	4.05**	3.73	0.92	1.55	5.90	1.13	0.28	0.47	1.79
Canadian (10)	2.78**	4.00	1.44	0.60	7.40	1.22	0.44	0.18	2.25
Canadian-Dutch	-0.16	-0.27	1.71	-4.31	3.77	-0.08	0.52	-1.31	1.14
Pooled (33)	4.25**	3.86	0.85	1.85	5.88	1.17	0.26	0.56	1.79

*Note.[a] Raw effect size units of measure are as follows: S24 raw unit=raw S24 score; SESAS=% score; PSI subscales=raw subscale scores (i.e., the number of items selected from each 20 item subscale). [b]df=31; [c]df=29. One pre-treatment questionnaire was not available. *p<0.01; **p<0.001.*

A post hoc analysis of the self-report data for Maintainers and Non-maintainers was undertaken to determine the extent to which the Non-Maintainers lost gains in attitudes, confidence, and perceptions. Means and standard deviations for the Dutch Maintainer and Non-Maintainer groups for the S24, SESAS, and the sub-scales of the PSI are also shown in Figures 2A-C. Improvements at F2 for Maintainers were 36.44% and 28.17% for the S24 and SESAS, and 43.87%, 66.87%, and 40.13% for the struggle, avoidance, and expectancy subscales of the PSI. Improvements at F2 for the Non-Maintainers were 23.81% and 17.93% for the S24 and SESAS, and 26.67%, 20%, and 38.18% for the struggle, avoidance, and expectancy subscales of the PSI. No analysis of the Canadian Maintainers and Non-Maintainers was performed because self-report data were available for only 1 of the Non-Maintainers.

SPQ

Table 4 presents the frequency and percentage of responses to the SPQ for the Dutch and Canadian groups at F2. For brevity, F1 responses are not presented; however, they are referred to for purposes of giving context to F2 responses when needed. Although the questionnaire return rate for the Canadian group was substantially less than that of the Dutch group, the results were generally similar across all items except for item 14. For example, although there was clearly a small proportion of Dutch participants who were not satisfied, the majority of Dutch and Canadian participants who returned questionnaires were satisfied with their speech immediately after therapy (item 2) and at F2 (item 3). It is notable that 2 of the Dutch participants who were very *dissatisfied* with their speech at F2 were maintaining 79.69% and 65.60% improvement in the beyond-clinic condition (and a mean of 58.10% and 65.60% improvement in the within-clinic conditions respectively). Interestingly, the proportion of Dutch and Canadian respondents who rated their current speech fluency as generally good at F2 (74% and 88% respectively; item 4) generally reflects the proportion of the groups who were Maintainers (71% and 86% respectively). It is notable that a larger proportion of the Dutch respondents (65%) than the Canadian respondents (44%) indicated that they had the skills to sound normal when controlling speech most of the time to almost always (item 7). However, overall the results were generally similar in that only a small proportion of each group reported that they seldom had the skills to sound normal when controlling speech.

Table 4. Frequency (proportion) of Dutch (n=23) and Canadian (n=10) group responses to the Speech Performance Questionnaire at 2 years follow-up.

Item			*Item*		
	Frequency (%)			*Frequency (%)*	
Response Option	*Dutch*	*Canadian*	*Response Option*	*Dutch*	*Canadian*
1. Satisfaction with speech before therapy			2. Satisfaction with speech immediately after therapy		
a. very satisfied	0	0	a. very satisfied	16 (69.6)	9 (100)
b. generally satisfied	4 (17.4)	0	b generally satisfied	6 (26.1)	0
c. generally dissatisfied	13 (56.5)	3 (33.3)	c. generally dissatisfied	0	0
d. very dissatisfied	6 (26.1)	5 (66.7)	d. very dissatisfied	1[a] (4.3)	0
3. Current rating of speech satisfaction			4. Current rating of speech fluency		
a. very satisfied	0	0	a. very good	0	0[c]
b. generally satisfied	15 (65.2)	7 (77.8)	b. generally good	17 (73.9)	7 (87.5)
c. generally dissatisfied	4 (17.4)	2 (22.2)	c. generally poor	5 (21.7)	0
d. very dissatisfied	4[b] (17.4)	0	d. very poor	1 (4.3)	1 (12.5)
5. Now have necessary skills to control speech			6. Now have necessary skills to sound fluent		
a. almost always	2 (8.7)	2 (22.2)	a. almost always	3 (13.0)	2 (22.2)
b. most of the time	11 (47.8)	3 (33.3)	b. most of the time	11 (47.8)	5 (55.6)
c. some of the time	7 (30.4)	4 (44.4)	c. some of the time	7 (30.4)	1 (11.1)
d. seldom	3 (13.0)	0	d. seldom	2 (8.7)	1 (11.1)
7. Now have necessary skills to sound normal when controlling speech			8. I use my speech controls		
a. almost always	3 (13.0)	2 (22.2)	a. almost always	0	0
b. most of the time	12 (52.2)	2 (22.2)	b. most of the time	4 (17.4)	2(22.2)
c. some of the time	6 (26.1)	4 (44.4)	c. some of the time	15 (65.2)	5 (55.6)
d. seldom	2 (8.7)	1 (11.1)	d. seldom	4 (17.4)	2 (22.2)

9. Now able to speak normally without thinking about controls			10. Now feel like a normal speaker		
a. almost always	3 (13.0)	1 (11.1)	a. almost always	1 (4.3)	0
b. most of the time	8 (34.8)	2 (22.2)	b. most of the time	11 (47.8)	3 (33.3)
c. some of the time	9 (39.1)	3 (33.3)	c. some of the time	8 (34.8)	5 (55.6)
d. seldom	3 (13.0)	3 (33.3)	d. seldom	3 (13.0)	1 (11.1)
11. As a result of therapy my speech fluency is			12. I found the Institute's therapy program to be		
a. much improved	8 (34.8)	3 (33.3)	a. very helpful	7 (30.4)	7 (77.8)
b. moderately improved	4 (17.4)	4 (44.4)	b. moderately helpful	12 (52.2)	2 (22.2)
c. slightly improved	8 (34.8)	2 (22.2)	c. slightly helpful	4 (17.4)	0
d. not improved	3 (13.0)	0	d. not helpful	0	0
13. Prefer stuttering over controlled speech			14. Currently consider myself a stutterer		
a. all of the time	0	1[d] (12.5)	a. yes	12[e] (52.2)	8[d] (100)
b. most of the time	4 (17.4)	0	b. no	11[f] (47.8)	0
c. some of the time	8 (34.8)	3 (37.5)			
d seldom	11 (47.8)	4 (50)			
15. Attribute speech improvements to[g]			16. In order to be fluent I must pay attention to my speech		
a. CSP	19[h] (90.5)	6[d] (75.0)	a. almost always	9 (39.1)	3[d] (37.5)
b. Other therapy	0	0	b. most of the time	10 (43.5)	4 (50.0)
c. Other factors	1 (4.8)	0	c. some of the time	4 (17.4)	1 (12.5)
d. a and c chosen	0	1 (12.5)	d. seldom	0	0
e. a and b chosen	1[i] (4.8)	0			
f. a, b, and c	0	1 (12.5)			
17. My fluency skills "work"			18. As a result of ISTAR therapy confidence in my ability to speak is:		
a. almost always	1 (4.3)	1[d] (12.5)	a. much improved	8 (34.8)	4[d] (50.0)
b. most of the time	14 (60.9)	3 (37.5)	b. moderately improved	10 (43.5)	3 (37.5)
c. some of the time	6 (26.1)	4 (50.0)	c. slightly improved	4 (17.4)	0
d. seldom	2 (8.7)	0	d. not improved	1 (4.3)	1 (12.5)
19. As a result of ISTAR therapy general confidence is:					
a. much improved	8 (34.78)	4[a] (50.0)			
b. moderately improved	8 (34.78)	2 (25.0)			
c. slightly improved	7 (30.43)	1 (12.5)			
d. not improved	0	1 (12.5)			

Note. [a]*This participant wrote: "Directly after therapy it went very bad at home, later on it got better. I had to get used to the transfer to my home."* [b]*In spite of being very dissatisfied, one participant wrote "I do have the feeling that I speak better than before therapy (and the possibilities to do so)."* [c]*n=8: one participant chose both b and c.* [d]*n=8: one participant did not complete the items 13 to 19 on the back page of the questionnaire at F2.* [e]*n=23; one participant said "only in the 3rd place. Stuttering is a characteristic, a part of my personality."* [f]*One participant wrote "I think stuttering is a terrible word." Another participant wrote "I never thought like this."* [g]*Responses d, e, and f reflect combinations of responses made by participants.* [h]*n=21; one participant did not complete this item; another participant did not choose from the offered responses but wrote "especially the self confidence is strongly increased."* [i]*This participant indicated that personal growth also contributed to improvement.*

There were notable similarities in the proportions of the Dutch and Canadian groups who preferred stuttering over controlled speech some of the time (35% and 38%) and those who seldom preferred to stutter (48% and 50%) (item 13). As well, there was a notable similarity in the proportions of the Dutch and Canadian groups who indicated that they needed to pay attention to speech almost always to be fluent (item 16): 39% and 38% respectively.

The most surprising result was the difference between the groups on item 14. This item asked participants to indicate whether or not they presently considered themselves to be a stutterer. At F2, 11 (48%) of the Dutch group indicated that they did not consider themselves a stutterer. Five of these participants showed a change over time. That is, at F1 they considered themselves as a stutterer but at F2 they did not. These results are in contrast to zero participants in the Canadian group in this study and in contrast to 2 participants in Boberg and Kully (1994) who did not consider themselves a stutterer at follow-up. A post hoc analysis of %SS at F2 revealed no significant difference between those who did and did not perceive themselves to be a stutterer (Mann-Whitney U=36, C=29, p>0.05).

Summary

Stuttering frequency measures indicated that both groups made substantial reductions in %SS at F2 compared to Pre and that all Pre-F2 contrasts for the groups individually and for the pooled %SS data were statistically significant. The %SS d effect sizes were medium or typical (0.52) for the Dutch group, larger than typical (0.86) for the Canadian group, and typical to larger than typical (.69) for the global treatment effect. Seventy-one percent of the Dutch group as compared to 86% of the Canadian group were classified as maintaining clinically meaningful reductions in stuttering frequency at F2. Speech rate outcomes were different for the groups. In contrast to the Canadian group, the Dutch group's mean SPM in the beyond clinic measure at F2 was lower than at Pre. However, the mean NAT rating for the Dutch group in the within-clinic monologue was essentially within 1 scale value of that previously reported for non-stuttering speakers on the Franken et al. (1992) scale. The mean Canadian NAT rating was within the range previously reported for non-stuttering speakers on the Martin et al. (1984) scale.

For the S24, SESAS and PSI subscales, statistically significant Pre-F2 differences were found for the groups individually and for the pooled data. The d effect sizes were all much larger than typical. Perceptions of speech performance measured by the SPQ were generally similar across the Dutch and Canadian groups except for perceptions of self as a stutterer (item 14). At F2, almost half of the Dutch group indicated that they did not consider themselves a stutterer. Five of these participants had changed from indicating that they perceived themselves as a stutterer at F1 to not doing so at F2. Interestingly, there was no significant difference in %SS between those who did and did not consider themselves a stutterer at F2.

Discussion

The primary purpose of this investigation was to evaluate the long-term effectiveness of the CSP within and across the Dutch and Canadian groups. Results indicate that Pre-F2 differences for stuttering frequency and self-report measures were statistically significant within and across the groups and that the large majority of participants in both groups were maintaining clinically meaningful reductions in stuttering at 2 years post-treatment. A secondary purpose of this investigation was to examine any differences in outcome and consider whether or not they may be due to cultural, methodological, or other factors. Two differences emerged: firstly, in the mean speech rates in the beyond-clinic measures of the two groups, and secondly, in the SPQ item that assessed perceptions of self as a stutterer.

Speech and self-report findings suggest that both Dutch and Canadian groups were maintaining treatment gains in (a) stuttering reductions and (b) improvements of speech related attitudes, confidence, and perceptions of struggle, avoidance, and expectancy to stutter at 2 years post-treatment. The global treatment effects for %SS and these self-report measures were positive: effect sizes ranged from typical to larger and much larger than typical in the behavioural sciences. The finding that the large majority of participants in both groups were categorized as Maintainers at F2 suggests that the stuttering frequency results also have clinical meaningfulness. NAT ratings for both groups also gave evidence of clinical meaningfulness of the %SS data. That is, reductions in stuttering frequency did not appear to be at the cost of relatively natural speech.

Although treatment effects for %SS were not statistically different between the groups the magnitude of the Dutch standardized effect size was somewhat lower than that of the Canadian group. In part, the higher Canadian standardized effect size was due to the smaller Canadian sample size; however, it is possible that other factors also may account for these results. Firstly, previous treatment history may have had an influence. In contrast to the Canadian group, for whom a minority (25%) had participated in two types of therapy programs prior to the CSP, the majority of Dutch participants (56%) had participated in two or more different types of therapy programs prior to the CSP. It may be that the Dutch group was comprised of more clients for whom stuttering appears to be intractable than the Canadian group. The diminished immediate post-treatment response (signalled by their unusually high Post %SS scores) of three of the Dutch participants lends support to this possibility. Secondly, differences between the groups in the proportion of clients who stuttered severely may be a factor. The proportion of participants who stuttered severely (*i.e.*, > 20 %SS) was greater in the Dutch group (25%; 6 of 24 participants) than the Canadian group (14%; 2 of 14 participants). Finally, differences in understanding due to translation from Dutch to English may have had an impact on clinicians and hence clients in the training of fluency and self-monitoring skills. Although it appears that the essential aspects of the speech and behaviour change concepts were understood by clients and clinicians, there may have been subtleties that were not conveyed that were discussed in the introductory cognitive-behavioural seminars with the clients and in the daily training sessions with the clinicians that were conducted in English.

It is encouraging that only two substantive differences between the Dutch and Canadian outcomes emerged in this study. In terms of the lack of a significant Pre-F2 difference in speech rate in the Dutch group's beyond-clinic measure, it seems unlikely that this would be due to differences in telephone practices between the two cultures given that expected increases in speech rate were achieved at post-treatment. However, the differences in the perception of self as a stutterer on the SPQ finding begs the question of whether it may be due to cultural differences in attitudes toward self, particularly because almost half of the Dutch group moved from having perceived themselves as a stutterer at F1 to not doing so at F2. This response is most unusual and is quite different from that of the Canadian group and past CSP outcome results.

The lack of a statistically significant difference in stuttering frequency between the Dutch participants who did and did not perceive themselves as a stutterer suggests that such self-perceptions are not likely based solely on stuttering frequency. One could posit that the difference may be due to translation of that item from English to Dutch. However, this is not likely the case since translation was done by native Dutch speakers who are also proficient in English. As well, post hoc back-translation (see Canino & Bravo, 1994) of the item indicates that the Dutch version of the item is linguistically equivalent to the English version. Despite linguistic equivalence, it is possible that the observed difference may be accounted for by "differences in the connotative meaning" of the item (Davidson, 1979, p. 142) or lack of conceptual equivalence (*e.g.*, see Canino & Bravo, 1994; Lonner, 1979). However, the concept of acceptance of one's self as a person who stutters some of the time but not all of the time was central to discussions of self-acceptance in both groups. Indeed, other factors may account for the difference. For example, as indicated in note "f" to Table 4, for at least one Dutch participant the difference may be due to an adverse reaction to the word "stutterer." For another participant, there simply was not the tendency to think in these terms. If differences in self-perceptions are replicated, then the next step of conducting theory-driven research investigating perceptions of and attitudes toward self as a stutterer within and across the groups will be warranted. If true cultural differences are found, then there will be implications for treatment and/or outcome measurement.

Taken together, the findings within and across the Dutch and Canadian groups in this study give evidence that the CSP appears to be effective in achieving durable and clinically important long-term reductions in stuttering frequency and improvements in associated speech attitudes, confidence and perceptions of struggle, avoidance, and expectancy to stutter. The results of this study were predictable given that there is a history of successful behavioural treatment for stuttering in the Netherlands (Franken et al., 1992, 1997). Thus, these results suggest that the CSP is sufficiently sensitive to the culture of Dutch adults who stutter and therefore, with the exception of the minor adjustment of the light touch fluency skill needed to preserve naturalness across Dutch dialects, adaptations to the treatment goals and processes do not appear necessary.

The finding that between 14 and 28% of the Canadian and Dutch participants respectively did not maintain clinically meaningful speech gains is generally in accord with previous CSP outcomes (Boberg & Kully, 1994) and that of Franken et al. (1997). Franken et al. concluded that the Dutch adaptation of the PFSP was not successful for about 30% of clients. These results indicate that research attention should now be turned to identifying treatments or treatment adaptations that are optimal for this subgroup of clients. Part of that process will require an understanding of how those who do and do not maintain fluency differ across a number of dimensions. Although pre-treatment stuttering severity has been shown to be a consistent but weak predictor of relapse (Craig, 1998), it may be that a combination of speech and attitudinal variables (Craig, 1998; Guitar, 1976) or combinations of neuromotor, linguistic, and emotional-motivational systems (McClean, Tasko & Runyan, 2004) influence relapse or treatment success. Indeed, it is possible that the success of the Maintainers who stuttered severely at pre-treatment was a result of an interaction between improved attitudes and use of speech change techniques. Interestingly, the Dutch Non-Maintainers in this study did not lose all gains in the self-report domains (*i.e.*, they were maintaining between 18% and 38% improvement across the S24, SESAS, and subscales of the PSI).

A limitation in this study is the difference between the groups in questionnaire return rates. Although the return rate of 56% for the Canadian group is adequate for analysis and reporting purposes (Babbie, 1973, as cited in Schiavetti & Metz, 1997), it is possible that the Canadian data is subject to response bias. It is tempting to speculate that participants did not return questionnaires because they were experiencing substantial regression or relapse in stuttering frequency. However, a post hoc analysis of stuttering frequency between the Canadian participants who did and did not return questionnaires indicated no significant difference between the subgroups (Mann-Whitney U=10, C=8; p>0.05). Thus, stuttering frequency does not appear to be the major reason why Canadian participants did not return questionnaires. A more likely reason is that receipt of therapy was not conditional upon participation in research as it was for the Dutch group. It is very likely that the response rate of the Dutch group in this study was at least partially attributable to the contingency upon which therapy was received.

A second limitation relates to the inability to compare NAT ratings in this study. Scales used in each group were those that were traditionally used in each country. Use of a common scale across the groups was mitigated by the different purposes of the independent naturalness studies underway in each country from which the NAT data were drawn; a common metric is needed in future investigations.

A third limitation relates to the methodology used to assess clinically meaningful speech gains in this study. Although there was a high level of agreement between categorizations made on the Dutch within- and beyond-clinic measures, which lends support for the validity of the telephone calls as a meaningful context within which to measure speech outcomes, it remains that the methodology needs validation with stakeholders, particularly clients. As indicated previously, the methodology was a preliminary attempt to characterize maintenance of speech gains that takes into account the inherent variability of stuttering and pre-treatment stuttering severities rather than absolute %SS in follow-up.

A fourth limitation is the difference in housing arrangements for the two groups. As indicated earlier, it is not customary to have CSP clients housed in one facility. Although it appears that the housing variable did not affect the overall outcome results, housing arrangements will need to be standard across the groups in future investigations.

In this study, evaluation of the effectiveness of treatment for individual participants was limited to maintenance of speech gains as opposed to maintenance of a combination of speech and attitudinal gains. Categorizing the success of treatment outcome for individuals for research purposes is fraught with complexity; there is an urgent need for prospective research that investigates the definition and development of a model that considers input from clients, significant others, clinicians, and researchers (see Finn, 2003; Yaruss & Quesal, 2002; Bothe 2004). Historically, definitions of relapse for research purposes have been based on absolute %SS criteria: that is, relapse has been defined as stuttering that is greater than 2 %SS or stuttering that is $\geq$3 %SS at follow-up (*e.g.*, Andrews & Craig, 1988; Blood, 1995; Craig et al., 1996). In our group, 3 %SS or less was previously defined as satisfactory outcome and between 3.1 and 6 %SS was defined as marginally satisfactory (Boberg & Kully, 1994).

There are at least four problems with such criteria. Firstly, some clients enter therapy with low levels of overt stuttering and thus have truncated ranges of improvement. Secondly, some clients whose pre-treatment stuttering is severe are placed in the "relapse" category in follow-up when in fact they are maintaining markedly improved levels of fluency in comparison to their pre-treatment levels. The data presented for client 3 in Boberg and Kully is illustrative. Client 3 had 53.73 %SS at pre-treatment, 21.76% at 1 year follow-up, and 8.5 % at 2 years follow-up. Based on the criterion that greater than 6 %SS at follow-up was unsatisfactory, this client was placed in the unsatisfactory category. In fact he was showing 59.50% and 84.18% improvement at 1 and 2 years post-treatment respectively. There is little doubt that these levels of improvement would be viewed as clinically meaningful by many clients. Thirdly, these criteria do not consider the individual's or group's retention of attitudinal changes (*e.g.*, cognitive and affective changes) that appear associated with recovery or successful management of stuttering (*e.g.*, Anderson & Felsenfeld, 2003; Plexico, Manning & DiLollo, 2005). Fourthly, these criteria have not been validated with input from clients.

Predictably the 2 year post-treatment outcomes of the first application of the CSP in the Netherlands are positive and similar to those of the Canadian group. Results indicate that that there were no differences across the cultures in outcomes in so far as they were measured by stuttering frequency and measures of speech related attitudes, confidence, and perceptions of struggle, avoidance, and expectancy to stutter. Although these results suggest that the CSP is sensitive to the culture of Dutch adults who stutter, the finding of a difference between the groups in perceptions of self as a stutterer at F2 and the change in self-perceptions in the Dutch group is intriguing. If replicated, further research into possible cultural differences in perceptions of self as a stutterer will be warranted.

Author Notes

We would like to thank the following people who have been of great help in the preparation of this manuscript: Kevin Bird, for his assistance with the analysis of treatment effect sizes and confidence intervals; Rogier Donders, for his statistical advice; Shelli Wright, for her assistance in collecting naturalness ratings which were part of her masters degree project submitted to the Department of Speech Language Pathology at the University of Alberta; and reviewers who provided constructive comments and suggestions for the revision of an earlier version of this manuscript. We also thank Ruud Bökkerink for being a driving force in the introduction of the CSP in the Netherlands.

Parts of this manuscript were presented at the 4th World Congress on Fluency Disorders, Montreal, Canada, 2003 and at the 26th World Congress of the International Association of Logopedics and Phonetics, Brisbane, Australia, 2004. A preliminary report of 1 year follow-up data for 9 of the Canadian participants in this study was reported in Langevin and Boberg (1993).

References

Algina, J., & Keselman, H.J. (2003). Approximate confidence intervals for effect sizes. *Educational and Psychological Measurement, 63,* 537-553.

Anderson, T.K., & Felsenfeld, S. (2003). A thematic analysis of late recovery from stuttering. *American Journal of Speech-Language Pathology, 12*, 243-253.

Andrews, G., & Cutler, J. (1974). Stuttering therapy: The relations between changes in symptom level and attitudes. *Journal of Speech and Hearing Disorders*, *39*, 312-319.

Andrews, G., & Craig, A. (1988). Prediction of outcome after treatment for stuttering. *British Journal of Psychiatry, 153*, 236-240.

Babbie, E.R. (1973). *Survey research methods*. Belmont, CA: Wadsworth.

Battle, D.E. (1993). Introduction. In D. E. Battle (Ed.), *Communication disorders in multicultural populations* (pp. xvi – xxiv). Boston, Andover.

Berry, J.W., Poortinga, Y.H., Segall, M.H., & Dasen, P.R. (2002). *Cross-cultural psychology: Research and applications*. New York: Cambridge University Press.

Betz, M.A., & Levin, J.R. (1982). Coherent analysis-of-variance hypothesis-testing strategies: A general model. *Journal of Educational Statistics, 7,* 193-206.

Bird, K.D. (2002). Confidence intervals for effect sizes in analysis of variance. *Educational and Psychological Measurement, 62,* 197–226.

Bird, K.D., Hadzi-Pavlovic, D., & Isaac, A.P. (2000). PSY [Computer Software]. Sydney: School of Psychology, University of New South Wales.

Bird, K.D., & Hadzi-Pavlovic, D. (2005). Studentized maximum root procedures for coherent analyses of two-factor fixed-effects designs. *Psychological Methods, 10*, 352-366.

Blood, G.W. (1995). POWER(2): relapse management with adolescents who stutter. *Language, Speech, and Hearing Services in the Schools, 26*, 169-179.

Bloodstein, O. (1987). A Handbook on stuttering (4th ed.). Chicago, IL: National Easter Seal Society.

Boberg, E., & Kully, D. (1985). *Comprehensive stuttering program.* San Diego, CA: College-Hill Press.

Boberg, E., & Kully, D. (1994). Long-term results of an intensive treatment program for adults and adolescents who stutter. *Journal of Speech and Hearing Research, 37,* 1050-1059

Boberg, E., & Sawyer, L. (1977). The maintenance of fluency following intensive therapy. *Human Communication, 2*, 21-28.

Bond, C.F., Wiitala, W.L., & Richard, F.D. (2003). Meta-analysis of raw mean differences. *Psychological Methods, 8*, 406-418.

Bothe, A.K. (2004). Evidence-based, outcomes-focused decisions about stuttering treatment: Clinical recommendations in context. In A. K. Bothe (Ed.), *Evidence-based treatment of stuttering: empirical bases and clinical applications* (pp. 261-270). Mahwah, NJ: Lawrence Erlbaum Associates.

Campbell, D.T. (1964). Distinguishing differences of perception from failures of communication in cross-cultural studies. In F. Northrop & H. Livingston (Eds.), *Cross-cultural understandings: Epistemology in anthropology* (pp. 308-336).

Canino, G., & Bravo, M. (1994). The adaptation and testing of diagnostic and outcome measures for cross-cultural research. *International Review of Psychiatry*, 6, 281-286. New York: Harper & Row.

Cohen, J. (1988). *Statistical power analysis for the behavioural sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates.

Cooper, E.B., & Cooper, C.S. (1993). Fluency disorders. In D. E. Battle (Ed.), *Communication disorders in multicultural populations* (pp. 189 - 211). Boston, Andover.

Cooper, E.B., & White, J. (1991). Unpublished data.

Craig, A. (1998). Relapse following treatment for stuttering: A critical review and correlative data. *Journal of Fluency Disorders, 23*, 1-30.

Craig, A., Hancock, K., Chang, E., McCready, C., Shepley, A., McCaul, A., Costello, D., Harding, S., Kehren, R., Masel, C., & Reilly, K. (1996). A controlled clinical trial for stuttering in persons aged 9 to 14 years. *Journal of Speech and Hearing Research, 39*, 808-826.

Davidson, A.R. (1979). Culture and attitude structure and change. In A.J. Marsella, R.G. Tharp, & T.J. Ciborowski (Eds.), *Perspectives on cross-culture psychology,* (pp. 137-157). New York: Academic Press.

Draguns, J.G. (1997). Abnormal behavior patterns across cultures: implications for counseling and psychotherapy. *International Journal of Intercultural Relations, 21*, 213-248.

Finn, P. (2003). Evidence-based treatment of stuttering: II. Clinical significance of behavioral stuttering treatments. *Journal of Fluency Disorders, 28*, 209-218.

Finn, P., & Cordes, A.K. (1997). Multicultural identification and treatment of stuttering: a continuing need for research. *Journal of Fluency Disorders, 22*, 291-236.

Franken, M., Boves, L., & Peters, H.F.M. (1997). Evaluation of Dutch Precision Fluency-Shaping Program. In E.C. Healey & H.F.M. Peters (Eds.), *International Fluency Association, 2nd World Congress on Fluency Disorders: Proceedings* (pp. 303-307). San Francisco: Nijmegen University Press.

Franken, M., Boves, L., Peters, H.F.M., & Webster, R.L. (1992). Perceptual evaluation of the speech before and after fluency shaping stuttering therapy. *Journal of Fluency Disorders, 17,* 223-241.

Franken, M., Boves, L., Peters, H.F.M., & Webster, R.L. (1995). Perceptual rating instrument for speech evaluation of stuttering treatment. *Journal of Speech and Hearing Research, 38,* 280-288.

Forne-Wastlund, H. (2004). Swedish Comprehensive Stuttering Program-Evaluation two to four years after treatment. In, A. Packman, A. Meltzer, & H. F. M. Peters (Eds.), *Theory, research and therapy in fluency disorders, Proceedings of the fourth world congress on fluency disorders* (pp. 189-194). Nijmegen, the Netherlands: Nijmegen University Press.

Guitar, B. (1976). Pretreatment factors associated with the outcome of stuttering therapy. *Journal of Speech and Hearing Research, 19*, 590-600.

Hillis, J.W., & McHugh, J. (1998). Theoretical and pragmatic considerations for extra clinical generalization. In A.K. Cordes & R.J. Ingham (Eds.), *Treatment efficacy for stuttering: A search for empirical bases.* (pp. 243-292). San Diego, CA: Singular Publishing Group, Inc.

Huinck, W.J., & Peters, H.F.M. (2004). Effectiviteitonderzoek bij stottertherapieen. Een longitudinale studie naar therapeutische effecten van drie verschillende therapievormen. Nijmegen: Radboud University, unpublished manuscript.

Ingham, J.C., & Riley, G. (1998). Guidelines for documentation of treatment efficacy for young children who stutter. *Journal of Speech, Language, & Hearing Research, 41*, 753-770.

Isaac, K. (2002). *Speech pathology in cultural and linguistic diversity*. London: Whurr Publishers.

Kraemer, H.C., Morgan, G.A., Leech, N.L., Gliner, J.A., Vaske, J.J., & Harmon, R.J. (2003). Measures of clinical significance. *Journal of the American Academy of Child and Adolescent Psychiatry, 42*, 1524-1529.

Kully, D. (1986). *Counting guidelines.* Edmonton, Alberta: Institute for Stuttering Treatment & Research, unpublished manuscript.

Kully, D., & Langevin, M. (1999). Intensive treatment for stuttering adolescents. In R. F. Curlee (Ed.), *Stuttering and related disorders of fluency* (pp. 139-149). New York: Thieme Medical

Kully, D., Langevin M., & Lomheim, H. (in press). Intensive treatment of adolescents and adults who stutter. In E. Conture & R. F. Curlee (Eds.), *Stuttering and related disorders of fluency, 3rd Edition*. New York: Thieme Publishers, Inc.

Langevin, M., & Boberg, E. (1993). Results of an intensive stuttering therapy program. *Journal of Speech-Language Pathology and Audiology, 17* (3 & 4), 158-166.

Langevin, M., & Boberg, E. (1996). Results of intensive stuttering therapy with adults who clutter and stutter. *Journal of Fluency Disorders, 21*, 315-327.

Langevin, M., & Kully, D. (2003). Evidence-based treatment of stuttering: III. Evidence-based practice in a clinical setting. *Journal of Fluency Disorders, 28*, 219-236.

Langevin, M., & Kully, D. (2004). Longitudinal treatment outcome: Four case studies. In A. Packman, A. Meltzer, & H.F.M. Peters (Eds.), *Theory, research and therapy in fluency disorders, Proceedings of the fourth world congress on fluency disorders* (pp. 195-203). Nijmegen, the Netherlands; Nijmegen University Press.

Lincoln, M., Onslow, M., Lewis, C., & Wilson, L. (1996). A clinical trial of an operant treatment for school-age children who stutter. *American Journal of Speech-Language Pathology*, *5*, 73-85.

Lonner, W.J. (1979). Issues in cross-cultural psychology. In A.J. Marsella, R.G. Tharp, & T.J. Ciborowski (Eds.), *Perspectives on cross-culture psychology* (pp. 17–45). New York: Academic Press.

Mackey, L., Finn, P., & Ingham, R.J. (1997). The effect of speech dialect on speech naturalness ratings: A systematic replication of Martin, Haroldson, and Triden (1984). *Journal of Speech, Language, and Hearing Research, 40,* 349-360.

Martin, R.R., & Haroldson, S.K. (1992). Stuttering and speech naturalness: Audio and audiovisual judgments. *Journal of Speech and Hearing Research, 35,* 521-528.

Martin, R.R., Haroldson, S.K., & Triden, K.A. (1984). Stuttering and speech naturalness. *Journal of Speech and Hearing Disorders, 49,* 53-58.

McClean, M.D., Tasko, S.M., & Runyan, C.M. (2004). Orofacial movements associated with fluent speech in persons who stutter. *Journal of Speech, Language, and Hearing Research, 47*, 294-303.

O'Brian, S., Onslow, M., Cream, A., & Packman, A. (2003). The Camperdown Program: Outcomes of a new prolonged-speech treatment model. *Journal of Speech, Language, and Hearing Research, 46*, 933-946.

Ornstein, A., & Manning, W. (1985). Self-efficacy scaling by adult stutterers. *Journal of Communication Disorders*, *18*, 313-320.

Perkins, W.H. (1981). Measurement and maintenance of fluency. In E. Boberg (Ed.), *Maintenance of fluency* (pp.147-178). New York: Elsevier North Holland.

Pickering, M., & McAllister, L. (2000). A conceptual framework for linking and guiding domestic cross-cultural and international practice in speech-language pathology. *Advances in Speech-Language Pathology, 2,* 93-106.

Plexico, L., Manning, W.H., & DiLollo, A. (2005). A phenomenological understanding of successful stuttering management. *Journal of Fluency Disorders, 30*, 1-22.

Schiavetti, N., & Metz, D.E. (1997). *Evaluating research in communicative disorders* (3rd ed.). Needham Heights, MA: Allyn & Bacon.

Scollon, R., & Scollon, S.W. (2001). Intercultural communication: A discourse approach (2nd ed.). Malden, MA: Blackwell Publishers.

Shapiro, D.A., Abbink, M., Bortz, M., Bruna, A., Cook, F., Dhu, P., et al. (2004). A multinational investigation of stuttering intervention: assumptions, practices, and lessons. In A. Packman, A. Meltzer, & H.F.M. Peters (Eds.), *Theory, research and therapy in fluency disorders, Proceedings of the fourth world congress on fluency disorders* (pp. 123-136*).* Nijmegen, the Netherlands: Nijmegen University Press.

Taylor, O.L. (1986). Historical perspectives and conceptual framework. In O. L. Taylor (Ed.), *Treatment of communication disorders in culturally and linguistically diverse populations* (pp. 3-18). San Diego, CA: College-Hill Press.

Ulliana, L., & Ingham, R.J. (1984). Behavioral and nonbehavioral variables in the measurement of stutterers' communication attitudes. *Journal of Speech and Hearing Disorders, 49*, 83-93.

Van Riper, C. (1973). *The treatment of stuttering.* New Jersey: Prentice-Hall.

Van de Vijver, F., & Leung, K. (1997). Methods and data analysis of comparative research. In J.W. Berry, Y.H. Poortinga, & J. Pandey (Eds.), *Theory and method. Vol I of Handbook of cross-cultural psychology (2nd Ed.)* (pp. 257-300). Boston, MA: Alyn and Bacon.

Webster, R.L. (1974). *The Precision Fluency Shaping Program: speech reconstruction for stutterers.* Roanoke, Virgina: Communications Development Corporation, Ltd.

Woolf, G. (1967). The assessment of stuttering as struggle, avoidance and expectancy. *British Journal of Disorders of Communication*, *2*, 158-171.

Yaruss, J.S., & Quesal, R.W. (2002). Research-based stuttering therapy revisited. *Perspectives on Fluency and Fluency Disorders [Newsletter of Special Interest Division 4, American Speech-Language Hearing Association], 12*(2), 22-24.

Chapter Five

The validity of a simple outcome measure to assess stuttering therapy

Huinck, W.J., & Rietveld, T. (2006). The validity of a simple outcome measure to assess stuttering therapy. Accepted for publication in Folia Phoniatrica and Logopaedica.

Abstract

Objective: The validity of a simple and not time-consuming self-assessment scale (SA) was tested to establish progress after or within stuttering therapy. Method: the scores on the SA scale were related to (1) objective measures (percentage stuttered syllables: %SS and syllables per minute: SPM) and (2) (self-)evaluation tests (self-evaluation questionnaires and perceptual evaluations or judgements of disfluency, naturalness and comfort by naïve listeners). Patients: Data were collected on two groups of stutterers at four measurement times: Pre-therapy, post therapy, 12 months after therapy and 24 months after therapy. The first group attended the Comprehensive Stuttering Programme (CSP): an integrated program based on fluency shaping techniques and the second group participated in a Dutch group therapy: the Doetinchem Method (DM) that focuses on emotions and cognitions related to stuttering. Results: Results showed similar score patterns on the SA assessment, the self-evaluation questionnaires, the objective measures over time, and significant correlations between the SA scale and SPM, %SS, PSIs and judged fluency on the T1-T2 difference scores. Conclusion: We concluded that the validity of the SA measure was proved and therefore encourage the use of suchlike instrument when (stuttering) treatment efficacy is studied.

Introduction

Stuttering is a complex phenomenon with a variety of characteristics and with a large range of severity levels. There are numerous reviews of literature examining treatment outcomes in children and adults (*e.g.* Blood, 1993, 1998; Bloodstein, 1995; Boberg & Kully, 1994; Conture, 1996; Thomass & Howell, 2001). Stuttering severity is not only based on the number of disfluencies in a person's speech. It involves a complex mixture of problems in speech production, intelligibility, emotions and cognitions and all of these aspects need to be evaluated in an efficacy study of stuttering therapy. That is why various methods of assessing stuttering severity have been described and used (Conture, 1996; Conture, 1997; Conture & Guitar, 1993; Thomas & Howell, 2001;). Although many treatment approaches are quite successful, it is well known that the success of treatment often decays after some time (Hasbrouck & Lowry, 1989; Ladouceur & Auger, 1980; Wagaman, Miltenberger & Arndorfer, 1993). Thus, all studies which are used to assess the success of therapies (should) include a longitudinal component: the fluency related behaviour of the participants is measured at a number of points in time. The measurement procedure itself is not unproblematic: a large number of test batteries is available but many of them are rather time consuming, specifically when one tests at both the overt and covert levels of stuttering. Translated tests which are frequently used in the Netherlands are: (1) The *Perceptions of Stuttering Inventory (PSI)* of Woolf (Woolf, 1967), (2) The Brutten *Speech Situation Checklist (SSC)* of Brutten, 1975, Brutten & Janssen, 1981, (3) The *Modified Erickson (1969) scale (S24)*, c.f. Andrews & Cutler, 1974 and (4) *The Lanyon Stuttering Severity Scale (SS Scale) (Lanyon, 1967).* Beside these questionnaires the percentage of stuttered syllables and spoken syllables per minute is frequently used for treatment evaluation. These percentages are often regarded as the ultimate criterion variable, for the obvious reason that they are measures in which subjective judgements play a minor role. The disadvantage of this procedure is the fact that it is time-consuming too. Another criterion is the evaluation of stuttered speech by judges who are not related to the speaker or a therapy. In a study of Franken, Boves, Peters & Webster (1995) for example, 'naïve listeners' had to judge 14 characteristics of speech on bipolar seven-point scales. This procedure is very time consuming because it must be carried out in a randomized experimental setting and each speech sample needs to be listened to.

There is also a simple self-assessment scale, with Equal Appearing Intervals (EAI) characteristics, ranging, for instance, from 1 (bad) to 10 (excellent): SA (Self-Assessment). This measure appears to carry a good face-validity, as the patient is likely to combine in his judgement both aspects which matter in stutter therapy: frequency and severity of stuttering and the subjective feelings associated with one's own speech behaviour. It is difficult to imagine that a stutterer would assess his/her verbal behaviour as 'quite good' because of low stutter frequencies if at the same time he/she experiences high tension as a result of modifications in speech behaviour. Obviously, it would be a good thing if the latter, simple and not time-consuming self-assessment scale yields valid scores. Of course, we do not know which aspect is assigned more weight in the SA-scores: frequency and severity of stuttering or subjective feelings associated with higher levels of tension. In this study we aim to assess the validity of this self-assessment scale (SA) by relating its outcomes with those obtained with measurement instruments which cover 'objective' stuttering and fluency counts, the self-perception of stuttering as measured by the PSI (Woolf, 1967), which is a weighted cumulative index of a number of aspects of stuttering, see below, and the perceptual evaluation of the stutterers' speech by listeners.

It was decided to use eight criterion variables in our study:
(1a,b) Objective criteria: Fluency and speech rate were measured by (1a) Syllables Per Minute (SPM) and (1b) Percentage Stuttered Syllables (%SS) respectively.
(2a,b,c) Subjective criteria based on the self-evaluation of the stutter to measure a multi-faceted attitude towards speakers' own verbal behaviour, the Perceptions of Stuttering Inventory of Woolf (1967) (PSI, with the three subscales: 2a avoidance, 2b expectancy; 2c struggle).
(3a,b,c) Subjective criteria based on judgements of listeners. We used two bipolar scales of Franken et al. (1995) (3a: fluency and 3b: naturalness) to assess the perceptual evaluation of the participants' speech. In addition, listeners had to indicate the level of comfort they experienced when listening to the speech of the participants (3c).

As it is well known that stutterers form a heterogeneous group and that some therapies focus on emotional aspects of stuttering, while other therapies directly aim at fluency enhancing techniques, two groups of adult stutterers were used who opted for two different therapies: 1) the Comprehensive Stuttering Programme (CSP) and 2) the Doetinchem Method (DM). We refer to the section *'Therapy programs involved in the investigation'* for more details. In this study we measured the sensitivity of the above-described measures to the longitudinal changes in both the DM and CSP. In a recent study of O'Brian, Packman, Onslow & O'Brian (2004), the comparative reliability of the percentage stuttered syllables and the score on a 9-point severity scale (scored by experienced judges) was investigated. The primary aim of this study was to compare

%SS scores and severity ratings in terms of (a) their distribution for a stuttering population (b) their relative reliability and (c) the degree to which scores on one scale predict scores on the other. The researchers concluded that both scales are reliable and can largely be used interchangeably for the measurement of stuttering. In the present experiment however, we are specifically interested in the stuttering participants' own as measured by the SA scores.

Method

Participants

Two groups of stuttering persons (randomly selected from a data pool of an efficacy study of the Radboud University Nijmegen Huinck & Peters (2004) participated in this study: a) Thirteen participants (9 men and 4 women; mean age 32 years, age range 17-53) followed the Comprehensive Stuttering Programme (CSP) and b) thirteen participants (8 men and 5 women; mean age: 21.3 years, range 17-32) followed the Doetinchem Method (DM).

Inclusion criteria were: (1) a reported onset of stuttering before the age of six (2); no reported problems in motor development (3); no reported concurrent problems in speech and/or language development (4); no reported use of medication that could influence respiration, phonation or articulation (5); no reported psychiatric problems (6); no reported hearing problems. The participants were not randomly assigned to the two therapies, but were found not to have different baselines (see section 6.1). Before the start of treatment the clients were extensively informed about the study and invited to participate. The participants who were prepared to join the study subsequently signed an informed consent prior to the first therapy session. In order to make both groups more comparable, we randomly selected 13 subjects from the CSP program.

Design and procedure

The effects of two types of stuttering therapies on a number of scales were tested at four different measurement times: pretherapy (T1); post-therapy (T2); 12 months post therapy (T3) and 24 months post therapy (T4). Stuttering therapy was followed between T1 and T2. Participants were allowed to select their own therapy programme and were therefore not randomly assigned to a treatment programme, which makes comparison between the two treatment programmes not possible. The primary aim of our investigation was to study how different measures relate to an efficient and low cost variable such as a Self-Assessment (SA) scale.

The dependent speech measures (%SS and SPM), the SA, the self-report questionnaire PSI and the speech samples that were used for the judgement of naturalness and fluency were all obtained at each measurement session (T1, T2, T3 and T4). The measurement sessions were independent of the treatment program (see the guidelines for documentation of treatment efficacy by Ingham & Riley, 1998), and thus took place at a different location (Radboud University Medical Centre Nijmegen) than the therapy location (a small town near Nijmegen, the Netherlands). It consisted of a registration of the interview, monologue, reading, telephone call on video and audio-tape, filling in a set of self-report questionnaires and a speech motor task with the

Nijmegen Speech Motor Test (NSMT); the results of the latter were not used in this study.

Therapy programs involved in the investigation
Fluency enhancing treatment
The Comprehensive Stuttering Programme (CSP) is a group therapy for adolescents and adults (Kully & Langevin, 1999). In this integrated treatment programme, fluency-enhancing techniques (*e.g.*, easy breathing, appropriate phrasing, easy onset, soft contacts, and continuous airflow/blending) are taught within a framework of prolonged speech. The programme aims at the reduction of core stuttering and learned struggle behaviours. Initially, the speech is reduced to speech rates of approximately 40 syllables per minute but during therapy speech rates systematically increase to a near normal rate of 190 syllables per minute. Clients learn to manage residual stuttering through tension modification and traditional stuttering modification techniques.

There are three treatment phases: acquisition of fluency skills and cognitive-behavioural strategies, transfer of skills into non-clinical environments, and skill maintenance in the months and years following therapy. Preparation for maintenance already begins during the acquisition phase but is more emphasised in the transfer phase (see for a full description of the CSP Kully & Langevin, 1999).

Cognitive treatment
The 'Doetinchem Method' (DM) is an intensive group therapy (five sessions of four consecutive days) that focuses on the social perspective of stuttering (Schoenaker & Schoenaker, 1975; Van Alphen & Hoogerwerf, 1984). The social context of the stutterer is the starting-point for the therapy which particularly targets the emotional and cognitive components of stuttering. The therapy also aims at improving speech fluency. Although the DM deals with stuttering from a broad perspective, it is mainly directed at two aspects of stuttering: first, the reduction of negative factors that maintain the stuttering problem and second, the enhancement of speech fluency. Based on behavioural treatment principles the client learns to describe, practise and use specific speech skills. The integration of thoughts and emotions is encouraged allowing for the clients' individual motivation and individual differences. Clients learn to reduce postural tension and tension in the respiratory muscles during speech. Stutter-free speech is trained, although improvement of communication as a whole is the main goal of the therapy. Negative emotions associated with stuttering are reduced, thus boosting the client's self-esteem (Janssen, 1973; Janssen, 1994; Schoenaker & Schoenaker, 1995; Van Alphen & Hoogerwerf, 1984).

Therapy content inventory
In order to obtain more detailed information about the actual content of the treatments, clinicians were asked to fill in a so-called therapy card (after each treatment session). On this card, the actual time spent on each treatment goal had to be indicated. The treatment goals were divided in treatment interventions targeting motor control and those directed at emotions and cognitions. The average time spent on these core aspects was provided for the programmes separately. In line with the original therapy goals, as described above, CSP clinicians indicated that they spent

73.3% of the time to changing speech behaviour and 26.7% of the time to changing emotions and cognitions (related to stuttering). The DM clinicians assigned respectively 36.1% to speech behaviour and 64.9% to emotions and cognitions. Thus, the CSP had stronger focus on speech-motor control issues, whereas DM spent relatively more time on modifying emotional and cognitive aspects.

Outcome Measures

(1) Self-Assessment (SA)
Participants had to judge their own speech on a speech satisfaction scale. This was a ten point rating scale (from 1-10) with 1 as the worst possible judgment and 10 as the best judgment. This scale is in line with the Dutch grading system in which 1=very bad; 2=bad; 3=very strongly insufficient; 4=strongly insufficient; 5=insufficient; 6=sufficient; 7=more than sufficient; 8=good; 9=very good; 10=excellent. The experimenter asked the participant to assign a general speech satisfaction score. Thus, the score was probably based on different aspects of stuttering like for example stuttering severity, negative emotional and cognitive reactions, reactions of listeners.

(2) Objective speech measures
Each participant was video-recorded for three minutes of talking time in an interview, a reading, a telephone call to an unfamiliar person and a monologue task. All participants were asked the same set of questions which were related to stuttering. With respect to the monologue, participants chose a topic from a set of topic cards (*e.g.*, vacation, hobbies etc.). Topics sets differed at each measurement session. The recordings were made in a building not associated with treatment by an unfamiliar research assistant who was not associated with the therapy program.
2a) Speech rate was assessed by counting the number of realized syllables per minute (SPM) in the interview. *2b)* Stuttering was assessed by counting the number of stutters in the three minutes interview sample, and expressed as a percentage of syllables stuttered (%SS), on the basis of the following formula: total syllables stuttered/total syllables × 100% = SS%. This was calculated in each session (T1, T2, T3 and T4) and for each participant. All syllable and stutter counts were made using electronic button-press event recorders (Boberg & Kully, 1985, 1994) by three trained raters (see Boberg & Kully (1994) for a description of the training program for the raters). The dependent variables (SPM and %SS) used in this investigation were based on three minutes speech samples from the interview only. This speech task was chosen because the speech situation during an interview is considered a typical spontaneous speech situation.

The reliabilities of SPM and %SS were calculated over all four speech samples by randomly assigning (with stratification for therapy) the speech samples of each participant to the raters. The participants were randomly assigned to the raters (3) with the exception of one rater who counted all different measures (T1, T2, T3 and T4) for the individual participant. Inter- and intra-reliability were calculated with the intra-class correlation coefficient (Schrout & Fleiss, 1979), with raters as a fixed factor and the stutterers as a random factor (which is equivalent to Cronbach's alpha). For inter-rater reliability, 159 (16.70%) speech samples from a total of 952 (238 sessions × 4 speech tasks) samples were recounted by all three raters (see also Langevin & Boberg, 1993). For intra-rater reliability, 112 (11.76%) samples were recounted by the same three raters, about 6 months after the first assessment; For the intra-reliability 'Time' was a fixed factor. The intra-rater reliability was above 0.95 for all three raters on both SPM and %SS counts, indicating a very high relative consistency of each independent counter.

The inter-rater reliability for both treatments and on SPM and %SS was also above 0.95. This indicates that there was a high agreement in SPM and %SS counts between the three counters. These high intra- and inter-rater reliability scores correspond to the scores obtained by Langevin & Boberg, (1993) and Boberg & Kully, 1994).

(3) Self-report questionnaire

In each session (T1, T2, T3 and T4), participants completed the *Perceptions of Stuttering Inventory (PSI)* questionnaire of Woolf (1967) to evaluate changes in feelings and emotions that are often associated with stuttering. The PSI examines the stutterer's perception of the presence of struggle, avoidance and expectancy of stuttering in his communication. It comprises 60 items that represent parameters of struggle (PSI-s), avoidance (PSI-a) and expectancy (PSI-e). Woolf suggests the PSI can be used to help the stutterer view his problem more objectively, to develop treatment goals and to assess progress. The protocol of the test indicates that levels of severity should be distinguished by the following division: Scores below 7 are mild; scores between 8 and 11 are moderate; scores between 12 and 15 are moderate-to-severe and scores between 16 and 20 are severe. Thus, higher scores on this questionnaire indicate a more negative perception of stuttering.

(4) Listeners' evaluation of stuttered speech

The listeners' evaluations of disfluency and naturalness were examined using two scales of the perceptual rating instrument developed by Franken et al. (1995) and Franke, Boves, Peters (1997). This instrument consists of 14 speech characteristics that have to be valuated on a (bipolar) seven-point scale: high pitch-low pitch; quick-slow; slovenly-polished; expressive-flat; shrill-deep; soft-loud; melodious-monotonous; tense-relaxed; weak accentuation-strong accentuation; unpleasant-pleasant (here called 'judged listener's comfort'); slurred-precise; halting-fluent ('judged fluency'); weak-powerful; natural-unnatural ('judged naturalness'). Contrastive terms label the extremes (analogously to the Semantic Differential Scale developed by Osgood, Succi & Tannenbaum, 1957).

Forty-two untrained female listeners (first year logopaedics students) with a mean age of 20.2 years (age range 17-26 years) evaluated four speech samples of 45 seconds (collected at T1, T2, T3 and T4) of 39 participants. These participants were randomly selected from the database of the Dutch evaluation study (see Huinck & Peters (2004) for an extensive description of this study). The raters had no experience in assessing or treating stuttering and were therefore considered as essentially naïve with respect to evaluating speech samples in this population. The speech samples were presented in a randomised order, with the restriction that two samples of the same speaker had to be separated by at least three samples of other speakers. The experiment took place in a classroom session. All participants listened simultaneously to the speech samples through the loudspeaker. Two of the above-described scales were related to the SA, the objective scores and self- evaluation scores: Fluent-Halting and Natural-Unnatural. In an additional session a group of 20 listeners (10 males and 10 females) with a mean age of 27.2 (range 22-37) evaluated the 'comfort' they experienced when listening to these speech samples. All listeners had a high educational level. On each scale, the scores of the 42 judges were averaged. The reliability of this average was very high (for each scale above .95).

Analysis of the data

Because it was not possible to randomly assign the participants to the treatment programmes, it was tested if the two groups of participants were equivalent in stuttering behaviour, measured at the pre-treatment session with the scales at issue (i.e. equal baselines in %SS, SPM, SA, three subscales of PSI, and the three perceptual evaluation scales). To test this, all dependent variables were analysed with t ests for independent samples (SPSS 12.0). None of the t-tests showed a significant difference at the 5% level on any variable, even without correction for multiple comparisons.

Two types of validity measures were calculated:
(a) Planned contrasts between measures obtained at T1 and T2 and between measures obtained at T1 and T4, by analysis of variance (with 'measurement session' as fixed within-subject factor) carried out on each of the dependent variables measured in the two therapy groups separately. A positive effect on the first contrast can be regarded as a minimal requirement for success of a therapy, as it is well known that stuttering improves largely immediately after treatment (cf. Langevin & Boberg, 1993; Craig, 1998; Kully & Langevin, 1999). If an outcome variable is not able to detect this expected change in stuttering behaviour, one should probably not use it. The second contrast concerns the scores obtained between the pre-treatment stage (T1) and the final stage (T4), here 24 months after treatment. This is a crucial contrast, as it reflects the desired change in verbal behaviour after possible effects associated with the process (for instance temporary disappointment as a result of experienced regression some time after the end of therapy) have faded away.

(b) Pearson correlations between the SA and the external criterion variables were calculated on the difference scores of all outcome variables obtained at T1 and T2 and those obtained at T1 and T4 to determine to what extent the values of the variables are 'proportional' to each other. If two scores are linearly related, high correlations will be found. The correlations on the change scores between T1 and T2 indicate how short term treatment effects on these scores are related, the change scores between T1 and T4 indicate how long term treatment effects on these scores are related.

Results

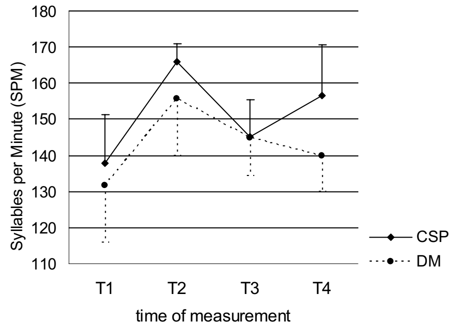
Analysis of Variance (repeated measures)
Analyses of variance (repeated measures), with 'measurement session' as within-subject factor were carried out on each of the outcome variables. In all but the SPM and PSIe, analyses the factor 'measurement session' was found significant ($p<0.05$). Planned contrasts (calculated between the reference value (scores obtained before therapy: T1) and T2 (immediately after therapy) and between T1 and T4 (24 months after therapy) for each of the two therapy groups separately are shown in Table 1.

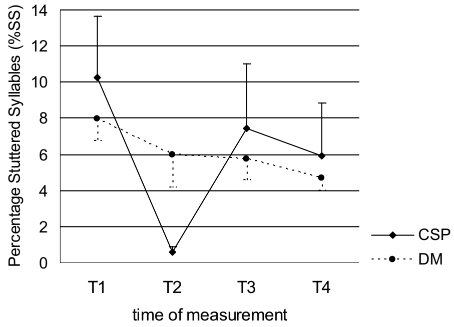
Table 1. Results of planned contrasts (p-values) for the percentage stuttered syllables (%SS), syllables per minute (SPM), the self-assessment scale (SA), three subscales of the Perceptions of Stuttering Inventory (PSIavoidance, PSIstruggle, PSIexpectancy) and judges of naïve listeners with respect to fluency, naturalness and comfort, expressed as positive (+) when a significant positive difference or a positive trend was found; negative (-) when a significant negative difference or a negative trend was found. CSP: n= 13, DM: n=13. Partial eta² is included as an indication of the effect size.

	CSP						*DM*					
	T1-T2			*T1-T4*			*T1-T2*			*T1-T4*		
	P		*Eta²*	*p*		*Eta²*	*p*		*Eta²*	*p*		*Eta²*
%SS	0.01	+	0.43	0.00	+	0.55	0.13	+	0.18	0.00	+	0.59
SPM	0.05	+	0.28	0.06	+	0.27	0.05	+	0.28	0.37	+	0.07
SA	0.00	+	0.77	0.00	+	0.68	0.00	+	0.75	0.00	+	0.77
PSI-s	0.02	+	0.73	0.03	+	0.57	0.04	+	0.76	0.06	+	0.60
PSI-a	0.17	+	0.59	0.15	+	0.51	0.16	+	0.55	0.09	+	0.48
PSI-e	0.00	+	0.25	0.02	+	0.28	0.01	+	0.30	0.03	+	0.41
Judged fluency	0.00	+	0.80	0.00	+	0.53	0.01	+	0.47	0.08	+	0.24
Judged naturalness	0.44	-	0.05	0.10	+	0.21	0.06	+	0.26	0.03	+	0.33
Listeners' comfort	0.89	-	0.00	0.02	+	0.40	0.02	+	0.36	0.05	+	0.29

Note. The eta-squared statistic describes the proportion of total variability attributable to a factor.
Although there is no significant effect (0) at T1-T4 measured with the PSIa and in dysfluency in the DM group, there are positive trends.

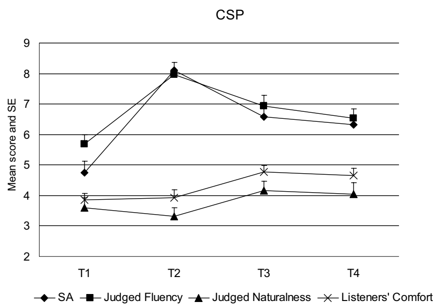
The table shows that gains scored on the Self Assessment scale (SA) largely agree with the gains obtained with scales which can be assumed to reflect stuttering more directly: %SS, PSI-s and Judged fluency. Significant improvements were observed immediately after therapy, with one exception: %SS in the DM group, but although there was no significant difference, the absolute values were in the expected direction (see Figures 1a and 1b). The improvements 24 months after therapy (T4) are less consistent over therapies. While scores obtained on most stutter-related variables (%SS, SA, PSI-s and judged fluency) showed significant long-term improvements in the CSP group, quite a large number of exceptions was found in the DM group: SPM, and Judged fluency, although there too, the non-significant tendencies were positive.

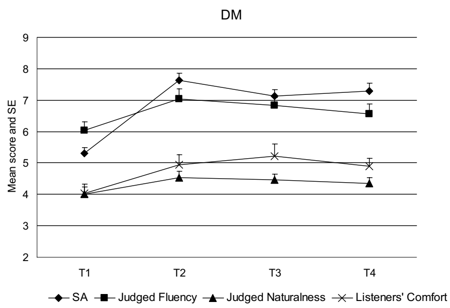




Figures 1a and 1b Mean (and SE) Syllables per Minute (figure1a) and mean (and SE) Percentage Stuttered Syllables (figure1b) in the CSP group (n=13) and DM group (n=13) at T1, T2, T3, and T4.

Judged naturalness and Listeners' comfort appeared to reflect a different dimension. In the CSP group Judged naturalness decreased (not significantly) from T1 to T2, it increased (not significantly though) from T1 to T4 (see Figures 2a and 2b). This was not unexpected, in light of the focus of that therapy. CSP focuses on the improvement of motor behaviour. The DM therapy which focuses on the social and emotional aspects of stuttering (see section 2) exhibited expected tendencies or significant improvements on Judged naturalness and Listeners' comfort, both from T1 to T2 and from T1 to T4. Apparently, scores on the Self Assessment scale strongly reflect measurements of disfluency (*e.g.*, %SS, SPM, PSI-s, judged fluency). On the other hand and not unexpectedly, scores on the Self Assessment scale do not reflect listener's appreciations of stutterers' speech in the CSP group. It is difficult to evaluate the contribution of the PSI scales. The PSi-s ('struggle') and PSI-a ('avoidance') scores are quite similar to the scores obtained with the less expensive ('expensive' in terms of time and effort) SA scale. PSI-e ('expectancy') showed a similar trend.





Figures 2a and 2b. Mean scores (the higher, the better) on the self-assessment scale (SA), judged fluency, judged naturalness and listeners' comfort in the CSP group (Figure 2a) and the DM group (Figure 2b) at T1, T2, T3, and T4.

Correlations between the outcome variables

Table 2 shows the correlation matrix on the difference scores T2-T1 and on the difference scores T4-T1. As can be seen in this Table, at the T2-T1 difference scores, the SA assessment significantly correlated with the SPM ($r= .42$; $p<0.05$), %SS($r=-.72$; $p<0.01$), PSIs ($r=-.60$; $p<0.01$); PSIa ($r=-.47$, $p<0.05$) and judged fluency $r=.43$; $p<0.05$). The correlation between judged naturalness and listeners' comfort was the strongest correlation. On the T1-T4 difference scores, there was no significant correlation with the SA assessment. All correlations at the T4-T1 difference scores were rather low (<0.3).

Table 2. Pearson correlations[a] (N=26)[b] between the SA scale and objective measures (%SS and SPM) and the (self-) evaluation tests PSIs, PSIa, PSIe, judgements of disfluency (Fluency), naturalness (Nat) and listeners' comfort (Comfort).

					T2–T1				
	SPM	*%SS*	*PSIs*	*PSIa*	*PSIe*	*SA*	*Comfort*	*Nat*	*Fluency*
SPM	1	-.610**	-.036	-.168	.115	.418*	.218	.222	-.352*
%SS	-.610**	1	.282	.259	-.103	-.716**	-.149	-.033	.503**
PSIs	-.036	.282	1	.648**	.215	-.595**	.243	.231	.045
PSIa	-.168	.259	.648**	1	.676**	-.470*	.296	.294	.178
PSIe	.115	-.103	.215	.676**	1	-.035	.062	-.085	.168
SA	.418*	-.716**	-.595**	-.470*	-.035	1	-.005	-.020	-.427*
Comfort	.218	-.149	.243	.296	.062	-.005	1	.812**	-.404*
Nat	.222	-.033	.231	.294	-.085	-.020	.812**	1	-.239
Fluency	-.352*	.503**	.045	.178	.168	-.427*	-.404*	-.239	1
					T4 – T1				
SPM	1	-.345*	.201	-.179	.204	.211	-.172	-.021	-.245
%SS	-.345*	1	-.036	.194	-.086	-.181	.121	.144	.218
PSIs	.201	-.036	1	.617**	.452*	.216	.081	.369	-.346
PSIa	-.179	.194	.617**	1	.693**	-.120	.160	.329	-.068
PSIe	.204	-.086	.452*	.693**	1	-.104	-.024	.054	-.142
SA	.211	-.181	.216	-.120	-.104	1	.018	.185	.050
Comfort	-.172	.121	.081	.160	-.024	.018	1	.440*	-.419*
Nat	-.021	.144	.369	.329	.054	.185	.440*	1	-.143
Fluency	-.245	.218	-.346	-.068	-.142	.050	-.419*	-.143	1

[a] **Correlation is significant at the 0.05 level (1-tailed)* and at the 0.01 level (1-tailed)**.*

[b] *Due to missing values in the PSI, correlations with the PSI-scales are based on 19 participants.*

Discussion

The aim of the present study was to investigate the validity of a simple and not time-consuming self-assessment scale by relating it to objective measures (%SS and SPM) and (self-)evaluation tests (including questionnaires and perceptual evaluations or judgements of disfluency, naturalness and listeners' comfort by naïve listeners). Results confirmed the validity of the SA-scale, as the scores followed the same pattern of results as the objective measure and the self-evaluation measures: a rather large improvement in speech behaviour and in (self-)evaluation, a marked regression at T3, but still, compared to T1, improvement at T4. When correlating the difference scores between T1 and T2 of the SA assessment with the other scores, *relatively high* correlations were found on the T2-T1 difference scores. The SA difference scores significantly correlated with measures which reflect overt stuttering behaviour: %SS, SPM, PSI-struggle and judged fluency. The only exception was the PSI-avoidance-scale. This latter measure is more related to covert features of stuttering. Hardly any other correlations were found for the time interval T4-T1. The only correlation with a relevant magnitude was the correlation between the PSIs and the PSIs (0.69).

As mentioned above, Pearson correlations were *relatively* high. The assessment of the significance and magnitude of Pearson correlations between the T2-T1 difference scores and the T4-T1 difference scores on the SA variable and the criterion variables might be negatively affected by a number of factors:
(1) Equal differences between scores obtained at T_i and T_{i+1} yield small amounts of variation and consequently low magnitudes of the correlations.
(2) SA-scores might not only be influenced by overt stuttering behaviour (*e.g.* fluency and speech rate), but also in case of disappointment when for example high expectations did not come true; this effect might be different for different (types of) stutterers and at different points in time
(3) If a measure has a low reliability, the reliability of the difference scores is limited to the square root of the product of the reliabilities at the two time points. As a result the correlation of two difference scores is limited to the square root of the reliability of the difference scores.
(4) Correlating difference scores may also yield quite ambiguous results as the value of the correlation depends on a large number of components. If we take for instance difference scores in SA and %SS, the ratios involved are: (SD post-SM)/(SD pre-SM) and (SD post-%SS)/(SD pre-%SS), and the correlations $SA_{T2}\%SS_{T2}$, $SA_{T1}\%SS_{T2}$, $SA_{T2}\%SS_{T1}$, $SA_{T1}\%SS_{T1}$, $SA_{T1}SA_{T2}$, $\%SS_{T1}\%SS_{T2}$ (cf. Gardner & Neufeld, 1987)[18]. As Gardner and Neufeld (p. 861) say "although the expression is (...) complex, there are some discernable patterns" (rephrased by us in our terms): (a) The effects of the test/retest correlations of the two measures (i.e., the values of rA2A1 and rB1B2) are such that, other things being equal, decreases in either term will lower the magnitude of the correlation between the change scores. (b) For given values of the test/retest correlations (i.e., rA2A1 and rB1B2) and of relative variabilities (i.e., M = sdB2/sdB1 and L = sdA2/sdA1), the magnitude of the correlation between the change scores depends upon a contrast between the correlations of A2 and A1 with B1 and of A2 and A1 with B1.
The low T4-T1 correlations are probably related to a differentiation between the participants in the score profiles two years after therapy. Long term measurements often show regression or relapse. Apparently, this relapse has different effects on

[18] *This can also be expressed as:*

$$r(A2-A1)(B2-B1)=\frac{M\times(L\times rA2B2)-(L\times rA2B1-rA1B1)}{\sqrt{(L^2+1-2\times L\times rA1A2)\times(M^2+1-2\times M\times rB1B2)}}$$

With: Two variables, A and B (for instance A=SA and B=%SS); measured at T1: A1 and B1 and T2: A2 and B2; r = Pearson correlation; sd = standard deviation; r(A2-A1)(B2-B1) = correlation of change scores in A and B. M = sdB2/sdB1 and L = sdA2/sdA1. Thus, we see that the correlation between the change scores is a function of 8 components: M, L, rA2B2, rA1B2, rA2B1, rA1B1, rA1A2, rB1B2.

individuals, some show more relapse in the fluency measures, whereas others show more relapse in measures of self-evaluations. This deviation in score profiles resulted in low correlation between the scores.
Thus, for the T1-T2 therapy outcome evaluation, the SA assessment seems to be a good and simple alternative for the default measurement of treatment efficacy. Moreover, the SA-scale is not only a valid score, but it also includes the most critical goal of stuttering treatment: self-judged- acceptability (cf. Ingham & Cordes, 1999).
In spite of the low correlations between T1-T4 difference scores, similar patterns of the scores over time at the group level were observed, except for the judged naturalness and listeners' comfort scores. Clearly, the latter two types of evaluation are measurements on a different dimension of speech. An instrument like the SA-scale is simple and not time-consuming, but might miss some subtle information. Particularly in studies with many participants, the SA-assessment might be a good evaluation measure to study fluency related changes over time. Assessment within (group or individual) therapy sessions may also benefit from the SA assessment can be used for interim evaluation of therapy progress. Although we might assume that scores on this scale are made up by a set of subjective features related to the stuttering problem (emotions, cognitions, attitudes, experiences of avoidance etc.), the overall SA-scale obviously does not reveal information on these separate features. For example, some stutterers may have high expectations of the therapy to which they have registered. When these expectations do not come true, feelings of disappointment might strongly decrease the SA-score. If researchers are interested in these specific features of the stuttering problem they have to recur to specific measures which focus on these aspects.

To conclude, results showed similar score patterns on the SA-scale, the self-evaluation questionnaires and the objective measures. As expected, judgements on the naturalness scale reflected the dissimilarity in the two treatment approaches. Relatively high correlations were found between the SA-scale and the fluency related measures like %SS, PSIs, Judged fluency and SPM. These results showed that the Self Assessment scale constitutes a valid instrument to assess stuttering, and therefore the use of this instrument should be encouraged when (stuttering) treatment efficacy is studied.

Acknowledgements

The research project was funded by the Dutch Health Insurance Board (CvZ, College van Zorgverzekeringen). The authors thank Esther Wouters for her assistance in the collection of the data.

References

Andrews, G. & Cutler, J. (1974). Stuttering therapy: The relation between changes in symptom level and attitudes. Journal of Speech and Hearing Disorders. *Journal of Speech and Hearing Disorders, 39,* 312-319.

Blood, G.W. (1993). Treatment efficacy in adults who stutter: Review and recommendations. *Journal of Fluency Disorders, 18,* 303-318.

Blood, G.W., Conture E.G. (1998). Outcomes measurement issues in fluency disorders. In C.M. Frattali (Ed.), *Measuring outcomes in speech-language pathology* (pp. 387-405). New York: Thieme.

Bloodstein, O. (1995). *A handbook on stuttering, Fifth edition.* San Diego: Singular Publishing Group, Inc.

Boberg, E. & Kully, D. (1985). *Comprehensive stuttering program.* San Diego, CA: College-Hill Press.

Boberg, E. & Kully, D. (1994). Long-term results of an intensive treatment program for adults and adolescents who stutter. *Journal of Speech and Hearing Research, 37,* 1050-1059.

Brutten, G.J. (1975). Stuttering: Topography, assessment and behavior change strategies. In J.Eisenson (Ed.), *Stuttering: A second symposium* (pp. 199-262). New York: Harper & Row.

Brutten, G.J. & Janssen, P. (1980). A normative and factor analysis study of the responses of Dutch and American stutterers to the Speech Situations Checklist (pp. 469-478). Washington DC: American Speech-Language-Hearing Association.

Conture, E.G. (1996). Treatment efficacy: Stuttering. *Journal of Speech and Hearing Research, 39,* 518-526.

Conture, E.G. (1997). Evaluating childhood stuttering. In R.F. Curlee, G.M. Siegel (eds.). *Nature and treatment of stuttering: new directions,* pp. 239-256. Needham Heights: MA, Allyn, Bacon.

Conture, E.G. & Guitar, B. (1993). Evaluating efficacy of treatment of stuttering. *Journal of Fluency Disorders, 18,* 253-287.

Craig, A. (1998). Relapse following treatment for stuttering: a critical review and correlative data. *Journal of Fluency Disorders, 23,* 1-30.

Franken, M.C., Boves, L., & Peters, H.F.M. (1997). Evaluation of Dutch Precision Fluency-shaping Program. In E. C. Healey & H. F. M. Peters (Eds.), (pp. 303-308). San Francisco: Nijmegen University Press.

Franken, M.C., Boves, L., Peters, H.F.M., & Webster, R.L. (1995). Perceptual rating instrument for speech evaluation of stuttering treatment. *Journal of Speech and Hearing Research, 38,* 280-288.

Gardner, R.C., & Neufeld, R.W.J. (1987). Use of the simple change score in correlation analyses. *Educational and Psychological Measurement*, *47*, 849-863.

Hasbrouck, J.M., Lowry, F. (1989). Elimination of stuttering and maintenance of fluency by means of airflow, tension reduction and discriminative stimulus control procedures. *Journal of Fluency Disorders, 14,* 165-183.

Huinck, W.J. & Peters, H.F.M. (2004). Effectiviteitonderzoek bij stottertherapieën. Een observationale studie naar therapeutische effecten van drie verschillende therapievormen.

Ingham, J.C. & Cordes, A.K. (1999). On watching a discipline shoot itself in the foot: Some observations on current trends in stuttering. In E.C.Healey & N. Bernstein Ratner (Eds.), *Stuttering research and practice: bridging the gap* (pp. 211-230). Mahwah: Lawrence Erlbaum Associates, Publishers.

Ingham, J.C. & Riley, G. (1998). Guidelines for documentation of treatment efficacy for young children who stutter. *Journal of Speech, Language, and Hearing Research, 41,* 753-770.

Janssen, P. (1973). *Vergelijking van effecten van twee stotterbehandelingen.* Praeventiefonds onderzoek nr. 28-33.

Janssen, P. (1994). *Helpt de Doetinchemse methode?* Stichting Stottertherapie Doetinchemse Methode.

Kully, D. & Langevin, M. (1999). Intensive treatment for stuttering adolescents. In R.F.Curlee (Ed.), *Stuttering and related disorders of fluency* (2nd ed., pp. 139-159). New York: Thieme Medical Publishers.

Ladouceur, R. & Auger, J. (1980). Where have all the follow ups gone? *Behavior Therapy, 3,* 10-11.

Langevin, M. & Boberg, E. (1993). Results of an intensive stuttering therapy program. *Journal of Speech-Language Pathology and Audiology, 17,* 158-166.

Lanyon, R.I. (1967). The measurement of stutter severity. *Journal of Speech and Hearing Research, 10,* 836-843.

O'Brian, S., Packman, A., Onslow, M., & O'Brian, N. (2004). Measurement of stuttering in adults: Comparison of stuttering-rate and severity-scaling methods. *Journal of Speech, Language, and Hearing Research, 47,* 1081-1087.

Osgood, C.E., Succi, G.J., & Tannenbaum, P.H. (1957). *The measurement of meaning.* Urbana: University of Illinois Press.

Schoenaker, T. & Schoenaker, T. (1975). *Stottertherapie.* Züntersbach: Rudolf Dreikers Institut für soziale gleichwertigkeit.

Shrout, P.E. & Fleiss, J.L. (1979). Intra class correlations: Uses in assessing rater reliability. *Psychological Bulletin, 2,* 420-428.

Thomas, C. & Howell, P. (2001). Assessing efficacy of stuttering treatments. *Journal of Fluency Disorders, 26,* 311-333.

Van Alphen, W. & Hoogerwerf, N. (1984). *Stotteren, helpt de Doetinchemse Methode? Een evaluatie-onderzoek naar resultaten van stottertherapie volgens de Doetinchemse Methode.* Stichting tot voorlichting over stotteren.

Wagaman, J.R., Miltenberger, R.G., & Arndorfer, R.E. (1993). Analysis of a symplified treatment for stuttering in children. *Journal of applied behavior analysis, 26,* 53-61.

Woolf, G. (1967). The assessment of stuttering as struggle, avoidance and expectancy. *British Journal of Disorders of Communication, 2,* 158-171.

Part II

Linguistic and motor factors in stuttering

Chapter Six

Gestural overlap in consonant clusters: Effects on the fluent speech of stuttering and non-stuttering subjects

Huinck, W.J., Van Lieshout, P.H.H.M., Peters, H.F.M. & Hulstijn, W. (2004).
Journal of Fluency Disorders, 29, 3-25.

Abstract

The present study was designed to investigate if persons who stutter differ from persons who do not stutter in the coproduction of different types of consonant clusters, as measured in the number of dysfluencies and incorrect speech productions, in speech reaction times and in word durations. Based on the Gestural Phonology Model of Browman and Goldstein, two types of consononant clusters were formed: homorganic and heterorganic clusters, both intra-syllabic (CVCC) and inter-syllabic (CVC#CVC). Overall, the results indicated that homorganic clusters elicited more incorrect speech productions and longer reaction times than the heterorganic clusters, but there was no difference between the homorganic and the heterorganic clusters in the word duration data. Persons who stutter showed a higher percentage dysfluencies and a higher percentage incorrect speech productions than PWNS but there were no main group effects in reaction times and word durations. However, there was a significant three-way interaction effect between group, cluster type and cluster place: homorganic clusters elicited longer reaction times than heterorganic clusters, but only in the inter-syllabic condition and only for persons who stutter. These results suggest that the production of two consonants with the same place of articulation across a syllable boundary puts higher demands on motor planning and/or initiation than producing the same cluster at the end of a syllable, in particular for PWS. The findings are discussed in light of current theories on speech motor control in stuttering.

Introduction

A challenging subject in the field of speech production research is the study of coordination between adjacent speech segments. In normal speech, segments tend to slide together because of a process called coarticulation, coproduction or blending. Coproduction is the overlap of speech segments or syllables and this overlap is often used to study different aspects of speech production. According to Tjaden (1999), the results from a large number of studies (*e.g.* Fowler, 1980; Fowler & Saltzman, 1993; Fujimura, 1986; Saltzman 1995; Salzman & Munhall, 1989) suggest that "a coproduction account of speech is well suited for explaining articulatory deficits in motor speech disorders because it conceptualizes speech as the variable overlap of articulatory gestures" (p. 604). One speech disorder that has been claimed to show problems in this respect is stuttering. Already in the sixties, Wingate (1964; 1969) described stuttering as a phonetic transition defect where, "…the difficulty is not manifested in the articulatory postures essential to that sound, but instead in moving on to the succeeding one(s)" (p.107). In recent years, a great number of studies have investigated the speech production skills of stuttering subjects (see for example Archibald & de Nil, 1999; Baken, McManus, & Cavallo, 1983; Bosshardt, 1999; Caruso, 1988; de Nil, 1995; Postma, Kolk & Povel, 1990; van Lieshout, Hulstijn & Peters, 1996a,b) and the findings indeed show that persons who stutter (PWS) differ from persons who do not stutter (PWNS) in the way they plan and/or execute speech gestures (i.e. as evidenced in a longer speech reaction time (RT) and/or word duration (WD), e.g. Peters, Hulstijn & Starkweather, 1989; van Lieshout, Hulstijn & Peters, 1996a; Diepstra, Huinck, Hulstijn & Peters 2001; Huinck, Wouters, Hulstijn, & Peters, 2001; but, see van Lieshout, Hulstijn, & Peters, 1996b). Howell, Au-Yeung and Sackin (2000) showed that consonant clusters at word initial position increase the chance of stuttering. However, clusters at other positions were not found to have a significant influence on stuttering frequency. Their findings can be related to a study by Santiago, MacKay, Palma and Rho (2000), who examined picture naming latencies for onset complexity and for number of syllables in normal speakers. It was found that latencies were longer for two syllable words than for one-syllable words and also longer for words beginning with consonant clusters than for words beginning with a single consonant. They found no interaction between onset complexity and vowel nucleus and coda complexity. However, in a re-analysis of these data, Roelofs (2002) suggested that the latency effect for cluster was actually a function of word length.

Clearly, cluster effects are important as they provide a window on the timing and sequencing of speech gestures (*e.g.*, Shaiman, 2001). Whether their influence is limited to initial word positions as argued in the papers discussed above (Howell et al, 2000; Santiago et al., 2000) is debatable. For example, a study by van Lieshout and colleagues (van Lieshout, Peters, & Bakker, 1997) showed that clusters had a consistent effect on reaction times and word durations for monosyllabic nonwords, regardless of their position (onset vs. coda). Given the fact that not much work is done on cluster complexity and stuttering, there is a clear need for further research in this area. To this end, the present study is designed to get more information about the influence of consonant clusters in coda and medial positions on speech reaction times and word durations of stuttering and non-stuttering subjects. Before discussing the usefulness of the coproduction account in stuttering research, we will first provide more general background information on this approach.

Coproduction has been modeled by Browman & Goldstein (1986; 1989; 1990a,b,c; 1992; 1997) in their Gestural Phonology Model. According to this articulatory-based phonological theory, phonological units, called gestures, are dynamical systems with a characteristic set of parameter values. Each action during the production of speech involves a synergy (the gesture) between various articulators to create a task-specific constriction of the vocal tract (the lips, jaw, tongue, glottis, velum), specified in terms of degree and location of the constriction. The implementation of these gestures is modeled in terms of task dynamics (Saltzman, 1986; Saltzman & Kelso, 1987; Saltzman & Munhall 1989). Crucial to this account is the notion of gestural co-production, that is, the way different autonomic gestures overlap in time and space to allow for a smooth integration of sequential movements within the domain of larger units of speech production like syllables and words.

Many researchers (*e.g.* Löfqvist and Yoshioka, 1981; Saltzman and Munhall, 1989; Löfqvist, 1990; Munhall & Löfqvist, 1992, Zsiga, 1994; Byrd & Tan, 1996; Tjaden & Weismer, 1998) have investigated the influence of speaking rate on gestural overlap. The assumption is that faster speaking rates cause gestures to slide together, resulting in a greater degree of overlap whereas at a slower speaking rate gestures tend to slide apart, resulting in less overlap (Tjaden & Weismer, 1998). Löfqvist and Yoshioka (1981) for example, found that a voiceless cluster of fricative and stop across a word boundary could be produced with one or with two laryngeal gestures depending on the speaking rate. At slow rates, two gestures were observed but at a fast rate only a single gesture occurred. In another experiment by Munhall and Löfqvist (1992), two speakers had to produce the utterance “Kiss Ted” at several speaking rates. Again, at fast rates, a single laryngeal gesture was found whereas at slow rates two separate gestures occurred. At an intermediate rate only one gesture was made, but this gesture showed traces of the two individual gestures.

Since a coproduction account of speech is claimed to be usefull for explaining articulatory deficits (*e.g.* Öhman, 1966; Tiffany, 1980; Young & Thompson, 1987; Löfqvist, 1990; Browman & Goldstein 1992; Liss & Weismer, 1993; Zsiga, 1994; Weismer, Tjaden & Kent 1995; Byrd, 1996; Tjaden, 1999; Tjaden 2000a,b; Zsiga, 2000; Nijland, Maassen, van der Meulen, Kraaimaat & Schreuder in press), studies have been carried out in which the amount of gestural overlap in different groups of speakers was measured. A common indirect way to measure gestural overlap is to calculate the formant transition slope[19] (*e.g.* Zsiga 1994; Tjaden, 1999). Tjaden (1999) for example, concluded that gestural overlap (measured in the F2 onset in consonant-vowel combinations) is reduced in speakers with neurological deficits who show scanning, or staccato-like speech patterns. In a subsequent study, Tjaden (2000b) investigated the effect of rate variation on coarticulation in people with Parkinson disease and in healthy control subjects. Formant transitions, expressed in the ratio of F2 onset frequency divided by F2 target frequency were used as a measure of coarticulation. The results showed that coarticulation increased with faster rates and decreased with slower rates. This relationship was more systematic for control speakers then in the people with Parkinson disease. Furthermore, data from a habitual speaking rate task suggested increased coarticulation for speakers with Parkinson relative to control speakers.

Formant transitions have also been examined in individuals who stutter, (*e.g.* Howell & Vause, 1986; Yaruss & Conture, 1993) and results showed that the coarticulation accompanying a dysfluency differs from the coarticulation of fluent speech (Robb & Blomgren, 1997). A deviating pattern of coarticulation in perceptually fluent speech of stuttering people is also found in other studies (*e.g.* Howell & Vause, 1986). This would suggest a potential timing and/or sequencing problem in PWS (see van Lieshout, 1995, for a review). However, Zebrowski, Conture & Cudahy (1985) could not confirm this.

Although these studies are of special importance in understanding articulatory organization of consonant-vowel (CV) and vowel-consonant (VC) transitions, little data has been published on how variations in the nature of consonant combinations that form clusters (CC) influence the amount of overlap within a cluster. Based on the Gestural Phonology Model (GPM) of Browman and Goldstein (1990c), specific predictions about the amount of overlap in different types of consonant clusters can be made. Browman and Goldstein organize gestural scores (a gestural score specifies the temporal activation intervals for each gesture in an utterance) into five articulatory tiers: velum, tongue tip, tongue body, lips and glottis. As a result, "gestures on different tiers may overlap in time and therefore proceed relatively independently of one another without perturbing each other's trajectories…" (p. 360). The authors give the example of the gestural representation of "must be" (/mʌst bi/) that will be pronounced as [mʌsbi] as a result of gestural overlap between the release of the consonants /s/ and

[19] *Formants are the local maximums in the vocal tract transfer function (Baken & Orlikoff, 2000) and can be measured in the spectrum of the acoustic signal. Usually second formant transition slopes are calculated in vowel-consonant (VC or CV) transitions to get information about the degree of temporal overlap between vowel and consonant.*

/b/, which effectively masks the release of the intermediate consonant /t/. In this example the amount of overlap was determined by the combination of two adjacent consonants (/s/ and /t/) that share the same place of articulation, which when combined with a bilabial consonant (such as /b/ in this example), results in a complete masking of the /t/ release.

According to this theory on articulatory organization, two types of consonant clusters are distinguished: homorganic and heterorganic consonant clusters. In homorganic clusters the consonants have the same place of articulation (i.e., belong to the same tier) whereas in heterorganic clusters the consonants have different places of articulation (*i.e.*, belong to different tiers). A homorganic consonant cluster is for example /xk/. Both /x/ and /k/ are velar consonants. The constriction for the realization of both consonants emerges from narrowing the space between tongue dorsum and velum. From a GPM point of view, the only difference between these consonants is the degree of constriction. Overlap between such gestures is restricted due to biomechanical and linguistic constraints in realizing a phonological contrast. On the other hand, in heterorganic consonant clusters, ample overlap between the gestures is possible. For example, in the heterorganic consonant cluster /ps/, /p/ and /s/ differ in place of articulation. That is, /p/ is a bilabial sound and /s/ an apico-alveolar sound. The gestural units involved in the production of this specific cluster will overlap considerably according to the GPM. This model would thus predict that heterorganic clusters (with gestures on different tiers) elicit shorter WD's than homorganic clusters (with gestures on the same tiers; see also Munhall, Fowler, Hawkins, & Saltzman, (1992) who measured gestural overlap in acoustic durations).

In (speech) motor research, RT is commonly used to index different stages in the planning and execution of motor processes (*e.g.* Sternberg, Monsell, Knoll & Wright, 1978). It is a general assumption that a more complex movement requires more time to prepare than a simple movement (see Schmit & Lee, 1999). Data from motor control experiments in limb movements, for example, suggest that complex movements elicit longer RT's because more time is needed for programming (and potentially initiating) these complex actions (*e.g.* Klapp, 1995, 1996; Henry, 1980). However, with respect to speech actions and consonant cluster complexity, it is unclear what kind of speech movements are more complex to produce: those belonging to homorganic consonant clusters or those belonging to heterorganic ones. It could be argued that the realization of heterorganic transitions is more complex because it involves the coordination of different articulators with potentially very different intrinsic biomechanical and physiological properties. Accomplishing such actions simultaneously might be a more complex task than doing one action after the other. This is in line with the assumptions underlying the Index of phonetic complexity (IPC), developed by Jakielski (2002). The IPC quantifies the phonetic complexity of words, based on eight factors. One of these factors is the consonant cluster type (homorganic versus heterorganic). According to the IPC heterorganic clusters are phonetically more complex than homorganic clusters. On the other hand, an old study by Coover (1923) indicated that transitions between keystrokes made by different fingers on the same hand are faster than transitions between repeated keystrokes made by the same finger (see also Rosenbaum, 1991 for review). Coover explained his findings in stating that in contrast to using different fingers, there is no overlap of movement possible during repeated keystrokes made by the same finger. Obviously, this mainly affects movement initiation and execution, but a recent study by Stoet and Hommel (1999) indicated a more general principle that involves action planning. According to these authors, action plans includes a temporal bond between selected action features. If two actions share the same effector system (that is, there is overlap between their action features), the overlap causes a delay in the preparation of a sequence of such actions compared to a situation where the overlap is reduced. Their theory found support in the planning and execution of finger movements. If we can extrapolate these findings on finger tasks to speech data, we could expect that the time to initiate a cluster produced by two different articulators (*e.g.* /pt/) is shorter than the initiation of a cluster produced with the same articulator (*e.g.* /st/) because in the former case, there is less overlap between action features Thus, we would predict longer RT's in the homorganic clusters than in the heterorganic clusters and, based on the GPM (more opportunity for temporal overlap in heterorganic than in homorganic clusters), shorter WD's in the heterorganic condition.

Apart from the type of consonant cluster (homorganic vs. heterorganic), it is known that syllable structure can play a role in the timing of articulatory gestures (Schiller, Meyer and Levelt, 1997). According to some researchers (*e.g.* Young and Thompson, 1987; Wheeldon and Levelt, 1995; Levelt and Wheeldon, 1994) the syllable is the basic unit for planning and producing speech. If there is a syllabic organization it is predicted that temporal adjustment and coarticulation within the syllable is stronger than between syllables (Browman and Goldstein, 1997; Levelt, Roelofs & Meyer, 1999). However, several recent studies on adults (Byrd, 1996; Schiller,1997) and children (Nijland et al., in press) could not find consistent evidence for gestural overlap differences between intra-syllabic and inter-syllabic consonant sequences.

Although the previous studies are not encouraging in terms of finding a consistent influence of syllable affiliation on coarticulation between consonants, there is a need for further testing of this issue as the syllabe unit is such a powerful concept in most theories on speech production (for an overview see Ziegler & Maassen, in press). Adhering to the notion that coarticulation within a syllable (CVCC) is stronger than between syllables (CVC#CVC) both conditions are included in the present study. If coarticulation within a syllable is stronger than between syllables, it is likely that the amount of gestural overlap between clusters within a syllable (*e.g.* CVCC) is more sensitive to cluster type variation (homorganic versus heterorganic cluster types) compared to the situation in which the cluster spans across the syllable boundary (*e.g.* CVC#CVC). Therefore, we expect to find a larger difference between homorganic and heterorganic clusters in clusters in coda position (CP) condition than when the cluster spans across the syllable boundary (SB).

In sum, the present study was designed to determine the influence of gestural overlap between two adjacent consonants in syllable code position (CP) and across the syllable boundary in bisyllabic words (SB) on speech RT and WD of PWS and PWNS for two types of clusters: homorganic and heterorganic. Based on the assumption that more time is needed for the planning and/or initiation of sequential movements from the same articulatory tier (Coover, 1923; Stoet & Hommel, 1999), it is predicted to find longer RT's in the homorganic condition compared to the heterorganic condition. Based on the Gestural Phonology model, which predicts more overlap and thus less execution time in heterorganic clusters, it is expected to find shorter WD's in the heterorganic condition. In addition, since PWS have been found to differ from PWNS in the time to complete speech production processes (*i.e.* longer speech RT's and WD's, e.g. Peters et al., 1989; van Lieshout et al., 1996a; Diepstra et al., 2001; Huinck et al., 2001), we expected that, if the above mentioned effects on RT's and WD's are found for normal speakers, PWS would show the same effects but to a stronger extent. That is, we expected a two-way interaction between group and type of cluster (homorganic vs. heterorganic).

Regarding the effects of cluster location, we expected to find a stronger WD effect of cluster type in the CP condition, based on the assumption that restrictions on coarticulation are less within a syllable than across a syllable boundary. If people who stutter have a speech timing and/or sequencing problem (see van Lieshout, 1995 for a review), this would obviously affect their capabilities to coarticulate gestures and as mentioned above, such a problem would be most visible in the production of different types of clusters in syllable coda position.

Method

Participants

Ten adult persons who stutter (PWS) and twelve adult persons who do not stutter (PWNS) participated in this study. The speech of twelve PWS was recorded and acoustically analyzed but because of the large number of missing values and extreme statistic outliers, the data of two participants in the PWS group were excluded from the statistical analysis (see also under "statistical analysis"). The two groups were matched as closely as possible on age, gender and educational level. Before onset of the experiment, participants had to fill in a questionnaire about their health and health-related topics (The "Stotteranamneselijst" which is a standard intake procedure for people who stutter at the University Hospital St Radboud, Nijmegen). All participants reported normal hearing acuity, normal language and voice quality and were native speakers of Dutch. Mean age of PWS was 23.9 years (SD=5.6; range=17-37) and for PWNS 24.1 years (SD=3.8; range=17-30). All participants were paid volunteers and gave their informed consent prior to participating in the experiment. Stutter severity was measured with the Stutter Severity Instrument (Riley, 1972) by an experienced clinician. The mean SSI score for all participants was 18.7 (SD=9.4). Individual data of the PWS-group are shown in Table 1, including information about the % dysfluencies and average speech rate per participant. This information is deemed important as it relates more directly than general SSI scores to potential speech motor control issues (van Lieshout et al., 1996b). Seven participants of the PWS-group were classified as mild, four as moderate and one as a severe stutterer (see Table 1). All PWS were measured before the onset of therapy and had not been in therapy for at least one year.

Table 1. The individual scores (and mean and SD) on the Stuttering Severity Index (SSI) and the mean number of syllables per minute (SPM) and mean percentage stutters (%SS) measured on the spontenous speech of each individual.

Participant	*Raw score SSI*	*SSI*	*Mean SPM*	*Mean %SS*
PWS 13	7	Mild	15	11
PWS 14	17	Mild	81	10
PWS 15	30	Moderate	12	6
PWS 17	8	Mild	17	2
PWS 18	8	Mild	15	6
PWS 20	27	Moderate	18	1
PWS 21	30	Severe	20	3
PWS 22	27	Moderate	12	9
PWS 23	13	Mild	17	4
PWS 24	20	Mild	88	18
Mean (*SD*)	18.7(*9.4*)			

Stimuli

The experimental material consisted of 12 mono- and 12 bi-syllabic pronounceable non-words with a constant onset and nucleus /ba:/. To exclude influences of higher order linguistic stages that precede the motor plan assembly stage (*e.g.*, van Lieshout et al., 1996a) only pronounceable non-words were selected. Each stimulus was presented two times in random order (this order was the same for each subject), adding to a total of 48 trials per subject. The monosyllabic trials were CVCC-sequences and the bi-syllabic trials were CVC#CVC-sequences[20].

Since manner of articulation influences the amount of overlap (Byrd, 1996), each consonant cluster in the set of stimuli was formed by a fricative and a plosive. To correct for possible consonant order effects (fricative-plosive versus plosive-fricative), we included each combination of consonants twice, once in fricative-plosive order and once in plosive-fricative order. With respect to the goals of this study, the stimuli varied for *Cluster Type*, with two levels: homorganic and heterorganic clusters; and *Cluster Location*, with two levels: mono-syllabic non-words with consonant clusters in coda position (CP) and bi-syllabic non-words with the consonant cluster spanning across a syllable boundary (SB). To match for transition complexity within the second syllable of a bi-syllabic string, the final consonant of the second syllable always had the same place of articulation as the first consonant of this syllable. A complete listing of all stimuli is given in Table 2. Cluster bigram frequencies[21] are given for each cluster (see also Table 2). Before the start of the experiment, participants were presented with 10 (different) exercise trials. These trials were not included in the analysis.

[20] *In Dutch, long vowels are spelled as single letters in open syllables and as geminates in closed syllables (in syllables in which the vowel is followed by at least one consonant). This means that each first syllable in the stimulus words has to be pronounced as a closed syllable. Therefore, Dutch speakers will always read the bi-syllabic words of the present study with the syllable boundary between the two consonants (e.g. /ba:s#tIt/ instead of /ba:#stit). See also Booij, 1995 for the phonology of Dutch.*

[21] *The term 'cluster frequency' indicates the frequency of occurrence of a specific pair of phonemes in Dutch. Cluster frequencies are based on 42 million tokens in the lexical database CELEX (Baayen, Piepenbroek and Gulikers, 1995).*

Instrumentation and procedure
Each participant was individually tested in the presence of two experimenters. Participants were seated at a distance of 1 m in front of a screen (15 inch) on which the stimuli were presented (font size 36). A second computer controlled the stimulus presentation using a delayed reading paradigm in which a warning signal (an audible click of 500 Hz during 1 ms) was immediately followed by the presentation of the stimulus word. The go-signal (a beep of 2000 Hz during 100 ms) was presented after a variable preparation period of 1500, 2000 or 2500 ms. The end of the response time (2000 ms after the go-signal) was indicated with a beep of 100 Hz during 100 ms. Following each trial there was a 1000 ms pause before the next trial started. The experimenter instructed the participant to respond as quickly as possible after the go-signal (but the participant received no feedback on acoustic reaction times). A microphone was placed approximately 30 cm in front of the participant's mouth. The audio signal was recorded on a DAT-recorder and also immediately read into a computer (sampling rate of 10KHz) using the Windaq recording program. Furthermore, the whole session was recorded on videotape. Besides acoustic data, physiological data on muscle activity, respiration and vocal fold vibration (cf., van Lieshout et. al., 1996a) were collected simultaneously, but these data will not be presented in this paper.

Data analysis

To exclude the influences of stutter events, only those utterances were analyzed that were judged as perceptually fluent utterances. To be accepted as fluent, an utterance had to satisfy two criteria. The first criterion was the absence of visual signs of struggle in the participant's face or body just before or during the token. One of the experimenters took note of these visual signs of dysfluency during the recording session. Second, the utterance could not contain audible hesitations, prolongations or repetitions of any kind. These acoustic signs of dysfluency were judged by two trained expert raters from an audio recording of the participant's speech. All stimuli on which judges disagreed were re-evaluated in an additional rating session (using the audio and/or video recordings). In this session the two raters discussed their original ratings until they agreed on all stimuli (see van Lieshout et al., 1996a for a similar procedure). In addition to dysfluencies, some utterances were (also based on consensus) classified as missing values (MV) because of for example inappropriate movements (*e.g.* swallowing) or momentary interruptions. Furthermore, using the same procedure, we classified reading errors (RE), slips of the tongue (SOT), too late reactions (LR) and too early reactions (ER). Trials were coded as LR when the participant was distracted and therefore started too late with the stimulus. Trials were coded as ER when the participant started before the go-signal. Acoustic speech reaction times (RT) and word durations (WD) were calculated from the remaining data set. RT is the time between the go-signal and the first acoustic response and WD is the time difference between the onset and the end of the acoustic response (*e.g.* van Lieshout, Starkweather, Hulstijn & Peters, 1995). Figure 1 shows WD and RT in an example of the acoustic signal for the stimulus /ba:x#kIk/. After a warning signal, the stimulus was presented during a variable preparation time interval (1500, 2000 or 2500 ms) followed by the go-signal, indicating to the participant to verbalize the nonword. The time of stimulus presentation was randomly assigned to the stimuli to keep the participant alert during the experiment. In doing so, it can be assumed that the motor plan for the stimulus is prepared before the presentation of the go-signal (c.f. Levelt, 1989; van Lieshout et al., 1996a,b; Roelofs, 2002a,b), and the reaction time only reflects motoric programming and initiation, that is, the adjustment of motor commands to the current phonetic context and the initiation of these commands (van Lieshout, 1995; van Lieshout et al., 1996a,b; van der Merwe, 1997).

Table 2. Experimental items systematically varied for 1) Cluster Type; 2) Cluster location and 3) Consonant Order. Cluster frequencies are given for each cluster according to the CELEX lexical database based 42 million lexical word forms (Baayen, Piepenbroek, Gulikers 1995).

Consonant transitions		*Cluster location*					
		Coda position (CP)			*Syllable Boundary Position (SB)*		
		Trial	*Phonetic transcription*	*Cluster frequency*	*Trial*	*Phonetic transcription*	*Cluster frequency*
Homorganic	f-p	BAAGK	/ba:xk/	0	BAAGKIK	/ba:xkIk/	216
	p-f	BAAKG	/ba:kx/	0	BAAKGIK	/ba:kxIk/	144
	f-p	BAAST	/ba:st/	3101	BAASTIT	/ba:stIt/	5079
	p-f	BAATS	/ba:ts/	4932	BAATSIT	/ba:tsIt/	3881
	f-p	BAAFP	/ba:fp/	0	BAAFPIB	/ba:fpIp/	54
	p-f	BAAPF	/ba:pf/	0	BAAPFIB	/ba:pfIp/	531
Heterorganic	f-p	BAAFK	/ba:fk/	0	BAAFKIK	/ba:fkIk/	165
	p-f	BAAKF	/ba:kf/	0	BAAKFIB	/ba:kfIp/	79
	f-p	BAASP	/ba:sp/	8	BAASPIB	/ba:spIp/	828
	p-f	BAAPS	/ba:ps/	17	BAAPSIT	/ba:psIt/	1078
	f-p	BAAGT	/ba:xt/	3225	BAAGTIT	/ba:xtIt/	6660
	p-f	BAATG	/ba:tx/	0	BAATGIK	/ba:txIk/	1064

In order to account for variations among participants in speech rate that might have influenced segment durations (*e.g.* Byrd & Tan, 1996; Tjaden & Weismer, 1998), WD's were also adjusted for the participant's overall speech rate relative to the other participants (see van Lieshout et al., 1995 for this procedure). These adjusted WD's were also included in the statistical analysis and the results were similar to the unadjusted results. Therefore, we only reported the unadjusted data in the results section.

To test for the reliability of our acoustic measures, ten percent of the data was randomly selected for re-analysis by a second person. Inter-rater reliability was measured with the intra-class correlation coefficient (ICC). For both RT and WD measures the ICC was > 0.99. There was no evidence of bias (mean difference < 1 ms for both RT and WD). In 90% of the cases the difference between the two raters was within 10 ms and in 95% of the cases the difference was within 20 ms.

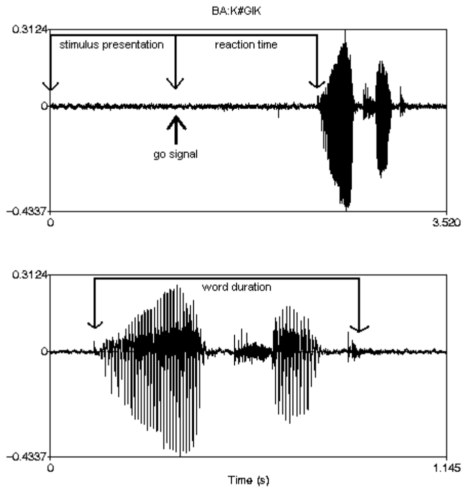


Figure 1. Example of the acoustic signal of the stimulus /ba:g#kIk/. The go-signal, reaction time (RT) and word duration (WD) are indicated with arrows.

Statistical analysis

Data were analyzed using a mixed analysis of variance design with repeated measures (GLM, SPSS 10.0) for each dependent variable with group (PWS vs. PWNS) as between-subject factor and two within-subject factors. The first within-subject factor c*luster type* had two levels: 1) homorganic consonant clusters (Ho) and 2) heterorganic consonant clusters (He) and the second within-subject factor c*luster location* had also two levels: 1) CP condition and 2) SB condition. The dependent variables were speech RT and WD. Before statistical analysis, data were averaged over the 6 repetitions within a stimulus category (Ho-CP; He-CP; Ho-SB; He-SB). Chi square (χ^2) tests were conducted on the number of incorrect speech productions to test for group differences, cluster type differences and cluster location differences. A stratified analysis procedure (Mantel-Haenszel) was used to test the difference of cluster type for PWS and PWNS. Mantel-Haenszel statistics can be used to test for independence between a dichotomous factor variable and a dichotomous response variable, conditional upon covariate patterns defined by one or more layer (control) variables. The Mantel-Haenszel common odds ratio is also computed for testing the homogeneity of the common odds ratio (SPSS, release 10, 2000).

Alpha was set at 0.05. One participant in the PWS group showed a large number of missing values (35.4%) and another participant in this group showed some extreme RT's and WD's (statistic outliers were 3 SD's above the group mean) resulting in a large range in the data set. Their data were excluded, leaving the data of 10 PWS in the data set for statistical comparisons.

Results

The results are presented in three sections. In the first section the incorrect speech productions are summarized. In the second section, the results of the speech RT data are presented and in the third section, the results of the WD data are presented.

Incorrect speech productions

Participants showed no signs (either verbally or non-verbally) of overall difficulties with the pronunciation of the stimuli during the experiment, attesting to the fact that the nonwords indeed adhered to normal phonotactic constraints for Dutch. However, sometimes participants did show incorrect speech productions. As mentioned in the Methods section, these incorrect speech productions were noted and classified as either slips of the tongue (SOT), reading errors (RE), late reactions (LR), early reactions (ER) and missing values (MV), and excluded for the statistical analysis.

Table 3 presents the sum of dysfluencies (DF) and incorrect speech productions subdivided in SOT, RE, LR, ER and MV for both PWS and PWNS. Stutter-like dysfluencies were not found for PWNS in contrast to PWS. Overall, PWS had a significantly [χ^2 =22.05, df=1, p<0.00] higher number of incorrect speech productions (13.3%) compared to PWNS (4.86%). As can be seen in Table 3, PWS showed a larger number (21) of slips of the tongue and early reactions (14) than dysfluencies (11). In general, the Ho condition elicited more incorrect speech productions (50) than the He condition (42), but this difference was not significant [χ^2 =0.651, df=1, p=0.42]. This cluster type influence was seen in both groups (Ho: PWS=33; PWNS=17; He: PWS=31; PWNS=11). Furthermore, there was no indication for a relation between the two variables as the Mantel-Haenszel test showed no significant dependency between group and cluster type [χ^2 =0.461, df=1, p=0.50].

Table 3. Number of dysfluencies (DF) and incorrect speech productions subdivided in slips of the tongue (SOT), reading errors (RE), late reactions (LR), early reactions (ER) and missing values (MV) for PWS and PWNS in both coda position (CP) and across the syllable boundary (SB). The total number of stimuli in the data set is 1056 (10 PWS and 12 PWNS each producing 48 stimuli).

Cluster type	*Cluster location*	*Subject type*	*Dysfluencies (DF)*	*Incorrect speech productions*					
				SOT	*RE*	*LR*	*ER*	*MV*	*Total*
Homorganic	CP	PWNS	0	2	4	1	0	2	9
		PWS	1	8	0	2	2	0	13
	SB	PWNS	0	1	3	1	1	2	8
		PWS	4	4	0	4	4	4	20
Heterorganic	CP	PWNS	0	2	1	0	0	2	5
		PWS	2	4	0	0	3	4	13
	SB	PWNS	0	2	1	1	1	1	6
		PWS	4	5	0	3	5	1	18
	Total	PWNS	0	7	9	3	2	7	28
		PWS	11	21	0	9	14	9	64

Reaction time (RT) data
Figure 2 presents the mean RT in of both groups in each condition. As can be seen in this figure, PWS on average were slower (mean=542 ms; SE=33 ms) than PWNS (mean=485 ms; SE=30 ms). However, our statistical analysis could not confirm this difference [$F(1,20)=1.895$, $p=0.184$]. The results yielded a statistically significant [$F(1,20)=8.992$, $p<0.01$] difference between Ho (mean=533 ms; SE=21 ms) and He non-words (mean=504 ms; SE=24 ms) and also between the SB condition (mean=544 ms; SE=24 ms) and the CP condition (mean=493 ms; SE=21ms), [$F(1,20)=25.397$, $p < 0.01$]. Obviously, the difference between the mono-syllabic CP condition and the bi-syllabic SB condition was due to a word length effect. Ho clusters elicited a longer RT than He clusters but this effect was largely based on the SB condition (56 ms) compared to the CP condition (2 ms), yielding a significant [$F(1,20)=9.446$, $p<0.01$] two-way interaction. The difference between Ho and He clusters was larger for PWS (mean difference=42 ms) than for PWNS (mean difference=15 ms) but this two-way interaction (group by cluster type) could not be confirmed statistically [$F(1,20)=1.992$, $p=0.173$]. However, the results did show a significant [$F(1,20)=6.19$, $p<0.05$] three-way interaction between group (PWS and PWNS), cluster type (Ho and He clusters) and cluster location (CP and SB condition); the difference between the two groups was largest in the Ho-SB condition (see Figure 2). In the PWNS-group, the difference between Ho and He in the CP condition corresponds to a similar difference in the SB condition but in the PWS-group this is not the case. In the SB condition, the RT for Ho stimuli is larger than for He stimuli and in the CP condition RT's for Ho are equal to or even smaller than RT's for He. This is not in line with our initial expectation to find a two-way interaction between group and type of cluster (homorganic vs. heterorganic) and a stronger WD effect of cluster type in the CP condition (see introduction).

Table 4 shows individual RT results in both Ho and He condition, showing individual differences between the participants. Most participants (75% of the PWNS-group and 90% of the PWS-group) showed a larger RT in the Ho condition compared to the He condition, regardless of the position of the cluster in the stimulus. This is in line with the more general assumption that it is more difficult to initiate Ho clusters compared to He clusters.

Word duration (WD)
Although the mean WD was longer for PWS (mean=613 ms; SE=23 ms) than for PWNS (mean=578 ms; SE=21 ms), our statistical analysis could not confirm this difference between the two groups [$F(1,20)=1.326$, $p=0.263$]. In contrast with the RT results, there was no effect of cluster type [$F(1,20)=0.254$, $p=0.620$] and, similar to the RT data, there was no interaction between cluster type and group [$F(1,20)=0.042$, $p=0.839$]. Furthermore, there was no significant [$F(1,20)=1.079$, $p=0.311$] three-way interaction between group, cluster type and cluster location (see Figure 3). As can be expected, due to different word lengths, there was a significant [$F(1,20)=102.560$, $p<0.01$] difference between the mono-syllabic CP condition (mean=549 ms; SE=15 ms) and the bi-syllabic SB condition (mean=643 ms; SE=17 ms).

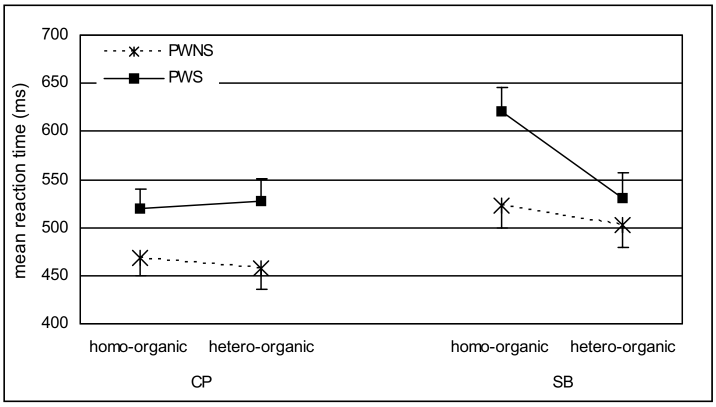


Figure 2. Mean reaction time (RT) in ms and standard error of PWS and PWNS for each cluster type (homorganic and heterorganic consonant clusters) in both cluster locations (CP and SB).

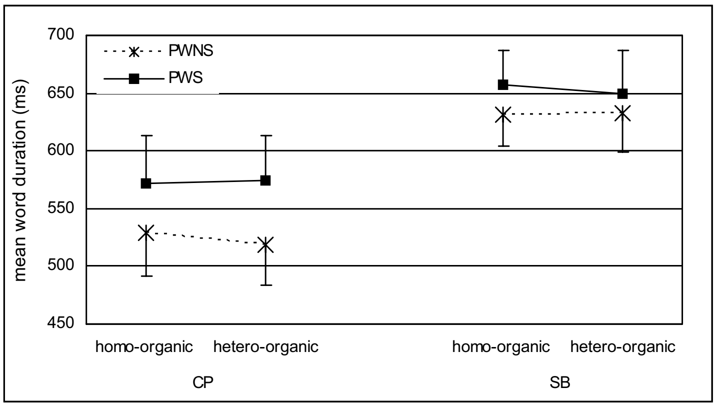


Figure 3. Mean word duration (WD) in ms and standard error of PWS and PWNS for each cluster type (homorganic and heterorganic consonant clusters) in both cluster locations (CP and SB).

Table 4 shows individual WD results in both Ho and He condition. Some participants spoke faster than others but most participants (83% of the PWNS-group and 70% of the PWS-group) showed a larger WD in the Ho condition compared to the He condition, regardless of the position of the cluster in the stimulus. This is in line with the more general assumption that it is more difficult to produce Ho clusters compared to He clusters.

Table 4. Mean reaction time (RT) and word duration (WD) (given in millisecond) of each person who stutters (PWS) and each person who does not stutter (PWNS), divided in Homorganic (Ho) and Heterorganic (He) results.

PWNS					*PWS*				
Subject	*RT (SE)*		*WD (SE)*		*Subject*	*RT (SE)*		*WD (SE)*	
	Ho	*He*	*Ho*	*He*		*Ho*	*He*	*Ho*	*He*
1	422(32)	413(33)	601(23)	571(24)	13	590(31)	529(32)	708(23)	667(23)
2	297(31)	290(32)	552(23)	541(23)	14	476(32)	459(32)	548(24)	522(24)
3	449(33)	493(32)	565(24)	556(23)	15	553(32)	550(31)	667(23)	611(23)
4	723(32)	745(31)	567(23)	552(23)	17	641(32)	613(31)	623(23)	601(23)
5	537(31)	509(31)	598(23)	630(23)	18	631(37)	684(38)	468(27)	453(28)
6	632(38)	552(32)	612(28)	608(24)	20	545(32)	512(32)	620(23)	558(24)
7	451(31)	448(31)	559(23)	572(23)	21	669(32)	526(33)	592(23)	601(24)
8	455(32)	413(32)	490(23)	478(23)	22	455(33)	397(34)	740(24)	709(25)
9	496(31)	493(31)	597(23)	590(23)	23	548(32)	436(31)	711(23)	744(23)
10	452(32)	400(32)	695(24)	717(23)	24	610(32)	597(33)	530(24)	549(24)
11	390(31)	358(31)	540(23)	528(23)					
12	653(32)	655(32)	586(23)	571(24)					

Participant 16 and 19 in the PWS group are excluded for the statistical analysis (see statistical analysis).

Discussion

This study was designed to investigate if stuttering participants differ from non-stuttering participants in the coarticulation of different types of consonant clusters at two different locations, namely coda position (CP) and across a syllable boundary (SB). Based on the Gestural Phonology Model of Browman and Goldstein (*e.g.* 1986; 1989; 1990a,b,c; 1992; 1997), two types of consonont clusters were formed: homorganic (Ho) and heterorganic (He) clusters. The main finding of this study is the significant three-way interaction between cluster type, cluster location and group. We found in the PWS-group longer RT's in homorganic clusters than in heterorganic clusters, but only in the SB condition and not in the CP condition (see Figure 2). In the PWNS-group, no such effect was found. Although we had originally expected that cluster type differences would be more pronounced in syllable coda position (where coarticulation effects are assumed to be stronger compared to coarticulation across syllable boundaries), the finding in itself is interesting as it clearly is in contrast with results of Howell et al. (2000), who only found effects of consonant cluster complexity in word-initial positions.

Main effects of Cluster type were only demonstrated in the RT data: longer latencies were found in the Ho condition than in the He data. This is in line with our expectations regarding the general effect of cluster type on RT, based on the early findings of Coover (1923) and the model proposed by Stoet and Hommel (1999). No overall group difference was found for RT and WD. This corresponds to findings reported by van Lieshout et al. (1996b), but in contrast with the RT results from Huinck et al. (2001) and Diepstra et al. (2001), who found longer RT's for PWS compared to control speakers (see also van Lieshout, 1995 for a review of RT studies in this area).

The number of dysfluencies were almost equal in both conditions (5 in the Ho condition and 6 in the He condition). There was also no effect of cluster type on the number of incorrect speech productions and on word duration. On average, PWS showed a higher percentage incorrect speech productions than PWNS (see Table 3).

In light of Browman and Goldsteins' (*e.g.* 1986; 1989; 1990a,b,c; 1992; 1997), ideas about gestural overlap, the results are somewhat ambiguous. Gestures on different tiers (as in the He condition in the present study) can overlap in time and, as a result, should elicit a shorter WD (execution time). This was not found. We did find some support for the claim by Stoet and Hommel (1999) that an overlap in action features between two successive actions causes a delay in response preparation, as evidenced in the longer RT for the homorganic clusters. These results are in contrast with the Index of Phonetic Complexity (Jakielski, 2002), although this theory is not specified in the context of RT paradigms but rather in terms of speech development. Our participants were all adults and a straightforward interpretation of the IPC under these conditions may not be justifiable.

That no effects of gestural overlap as a function of cluster type were found in the WD data might have to do with lengthening effects before the clusters (*e.g.* Fulk, 1999, who found evidence for lengthening before homorganic clusters in Old English) or with coda cluster reduction effects. Final lengthening effects appear (as in English) also in Dutch. This usually occurs in the final position of syllable or word. Because in both (CP and SB) conditions we are dealing with so-called "closed vowels", we expected to find this lengthening on the final consonant in the syllable which is in our stimuli always the final consonant of the cluster (either a plosive or a fricative). Fricatives are sensitive to final lengthening effects and therefore we tested for any effect of final lengthening in the WD data. We did find a significant ($p<0.05$) main effect of consonant order in the cluster (either plosive-fricative or fricative-plosive), but there was no interaction between consonant order and cluster location (CP and SB). Therefore, it seems we can exclude that final lengthening effects has a specific influence on our results.

We would like to raise some general points of discussion regarding the RT data. First, it can be argued that cluster frequency affects RT results. In Table 2 the cluster frequency (in Dutch) for each consonant sequence is given. In the CP condition most of the cluster frequencies are zero or near zero. This is shown in both Ho and He conditions. In the Ho condition there are two very frequent clusters (/st/ and /ts/) and in the He condition only one cluster was highly frequent (/xt/). If this had any influence on our data, it would have induced shorter RT's in the Ho-condition. However, this condition showed longer RT's. A possible effect of cluster type in the CP condition may have been attenuated as a result of the overall low cluster frequency. If these clusters were indeed fairly unfamiliar, a more automated pattern of coarticulation would have required some practice, especially for the He clusters, for which there is less natural restriction due to natural physical/physiological constraints. In the SB condition, the total frequency of the clusters was higher and more or less equal in both conditions (total cluster frequency of Ho=9905 and of He=9874). The more frequent a cluster is, the shorter the RT for that specific cluster. However, there were three very high frequent clusters: /ts/ and /st/ in the Ho-condition and /xt/ in the He-condition (cluster frequency above 1500). Since two out of the three clusters are of the Ho type, any influence of cluster frequency potentially would bias the Ho data towards faster RT's, the opposite of what we have found. In other words, if cluster frequency affected RT in this study, it would have *diminished* the cluster type effect. To test this assumption, we performed the statistical analysis with exclusion of the /ts/, /st/ and /xt/ stimuli for RT and WD with no difference in results. Thus, the RT data are not influenced by cluster frequency effects but seem to reflect an actual difference between Ho and He clusters.

The second remark relates to word length differences in the set of stimuli. The SB sequences consisted of two syllables, whereas the CP sequences consisted of one syllable. It has been shown that word length influences RT's in normal speakers (*e.g.* Jared & Seidenberg 1990; Klap, Anderson & Berrian 1973; Bachoud-Levi, Dupoux, Cohen & Mehler 1998; Erikson, Pollack & Montague 1970; van Lieshout et al., 1996a; 1996b) and PWS (*e.g.* Diepstra et al. 2001; Huinck et al. 2001; Peters et al. 1989; van Lieshout et al., 1996a). Longer utterances (as in the SB condition) are potentially motorically more complex (*e.g.*, Peters et al., 1989). That in itself does not need to trigger a significant group effect (van Lieshout et al., 1996b). However, if the demands on motor programming (and/or initiation) for bisyllabic words are already higher compared to monosyllabic words, and there is an additional demand on these stages due to the repetition of movements by the same articulator (homorganic cluster across the syllable boundary) as suggested by Stoet and Hommel (1999), the combined demands might push persons who stutter with potential limited speech motor skills to their limits. Hence, the three-way interaction for group, cluster type and location, with longer RT's for PWS in the two-syllable condition for homorganic clusters.

Reaction time may also be influenced (reduced) by the constant first syllable /ba:/ in the stimuli. Roelofs (2002b) tested the minimalist theory (articulation starts when the first speech segment has been planned) and the non-minimalist theory (larger units are planned and buffered before articulation is initiated) of speech planning by making a distinction between homogeneous and heterogeneous speech segments in bi-syllabic words (experiment 1) and in utterances consisting of two phonological words (experiment 2) in an auditory priming experiment. Homogeneous words in his experiments shared one or more speech segments, whereas heterogeneous words were different in form. The author found faster production latencies in the homogeneous condition compared to the heterogeneous condition. In the present study, all stimuli started (as in the homogeneous condition of Roelofs) with a constant onset- rhyme combination /ba:/ and it is therefore possible that this reduced the RT's. However, this possible reduction in RT would be equal for all our stimuli, and therefore it seems hard to argue in favour of a specific influence that would have biased the RT results. Also, by keeping the onset constant, possible systematic variations in RT and WD can be attributed with more certainty to non-initial sound influences. This is not possible in case the onset varies because uncontrollable interactions may occur.

Another issue that needs to be addressed in the context of our study is the use of perceptually fluent speech in PWS. This is a fairly common practice in most studies on speech motor behavior, based on the assumption that the speech motor deficits that are responsible for stuttering are ever present during speech production (Armson & Kalinowski, 1994). However, according to Armson and Kalinowsky (1994) the characteristics of these perceptually fluent speech samples may change as a function of a) context of the experimental samples; b) treatment history of participants; c) stuttering severity and d) developmental history of stuttering. They argue that results from research that has studied the perceptually fluent speech of PWS should be interpreted with caution. We acknowledge that for the present study this also could be a potential issue. However, in the absence of a significant group effect on RT or WD, it is unlikely that our sample is contaminated with a systematic bias on temporal aspects of speech production as related to either hidden disfluencies or compensatory behaviors in people who stutter. If such influences occur more randomly and in a subset of individuals (as can be expected), the chances of them underlying the three-way interaction in RT found in this study are remote. In addition, we do believe that for this type of research, which is based on a long-standing tradition in motor and linguistic studies for normal speaking participants, it is inevitable to use perceptually fluent speech samples. Including dysfluencies in the data set would contaminate the results and make a clear interpretation in terms of existing speech production models nearly impossible.

Our WD results do not indicate that gestural overlap is greater within syllables than across syllables (as suggested by Browman and Goldstein, 1997; Levelt et al., 1999 and as hypothesised in this study). In both the CP and SB condition, the WD difference between the two cluster types is almost equal in both groups of speakers. One can argue that there was no effect of WD found since differences in the constant onset rhyme combination /ba:/ may systematically compensate for differences in cluster duration. If such effects would have occurred, they should have influenced the word durations (especially in our monosyllabic stimuli). However, former studies (*e.g.* Shaiman, 2001; Hertrich & Ackerman 1995) showed that differences in WD do co-vary with differences in cluster duration. For example, Shaiman (2001) studied the kinematic and acoustic effects of coarticulation in singletons and clusters in coda position of a monosyllabic stimulus. Vowel reductions were never found to be strong enough to entirely compensate for such effects. Byrd (1996) also hypothesized that temporal coproduction in consonant sequences is greater within a syllable (intra-syllabic) than between syllables (inter-syllabic) but like us, she did not find consistent evidence of overlap differences between intra- and inter syllabic sequences (measured with electropalatography). She did find, on the other hand, that onset clusters are less overlapped than coda clusters and clusters that span across a syllable boundary. As mentioned in the introduction, it seems that the general assumption of a syllabic influence on coarticulation is hard to verify in a controlled experimental setting (see also Nijland et al., in press; Ziegler & Maassen, in press; Schiller, 1997).

The individual WD data of the present study showed, conform results of Byrd and Tan (1996), who reported individual differences in gestural overlap, that some participants spoke faster than others. However, most participants (83% of the PWNS-group and 70% of the PWS-group) showed a larger WD in the Ho condition compared to the He condition, regardless of the position of the cluster in the stimulus (see Table 4). This means that the individual WD data for the majority of our participants support the findings. Furthermore, since the WD results were similar to the results of unadjusted WD's (see the "data analysis" section), it can be assumed that the WD's were not influenced by the individual speech rate differences.

To conclude, effects of cluster type manipulation in the SB condition have shown to affect motor programming/initiation processes (reflected in RT) of fluent and dysfluent speakers differently. The Ho condition elicited longer RT's than the He condition. These results seem to agree with predictions that are based on the Gestural Phonology Model of Browman and Goldstein (1990c) and suggest that the production of two consonants that share their place of articulation across a syllable boundary demands more programming and/or initiation time than producing the same cluster within a syllable, particularly in PWS. Although this is a very interesting result, the data also raise some questions: Why did we not find a cluster type effect in the WD data? Why did we not find more dysfluencies and incorrect speech productions in the Ho condition compared to the He condition? What is the exact relation between consonant clusters and syllable boundaries? Further experiments are necessary in order to improve our understanding of the here presented effects of articulatory complexity of cluster sequences on the speech of stuttering participants. In these studies, onset clusters can be added, while controlling for word length. Furthermore, a simultaneous registration of RT and articulator movements, recorded with Electromagnetic midsagittal articulography (EMMA; van Lieshout & Moussa, 2000), might provide a better understanding of the relation between RT and gestural overlap.

Acknowledgements

This study was supported by the Dutch Health Care Insurance. The authors thank Judith Verberne and Esther Wouters for the assistance in the acoustic data analysis, Mariëtte Embrechts for her help in the scoring of the stutter severity and prof. dr. Rob Schreuder for his help in using CELEX.

References

Archibald, L. & de Nil, L.F. (1999). The Relationship between Stuttering Severity and Kinesthetic Acuity for Jaw movements in Adults who Stutter. *Journal of Fluency Disorders, 24,* 25-42.

Armson, J. & Kalinowski, J. (1994). Interpreting results of the fluent speech paradigm in stuttering research: Difficulties in separating cause from effect. *Journal of Speech and Hearing Research, 37,* 69-82.

Baayen, R.H., Piepenbroek, R., & Gulikers, L. (1995). The CELEX lexical database. [CD-rom]. Philadelphia, PA: Linguistic Data Consortium, University of Pennsylvania.

Bachoud-Levi, A.C., Dupoux, E., Cohen, L., & Mehler, J. (1998). Where is the Length effect? A cross-linguistic study of speech production. *Journal of Memory and Language, 39,* 331-346.

Baken, R.J., McManus, D., & Cavallo, S. (1983). Prephonatory chest wall posturing in stutterers. *Journal of Speech and Hearing Disorders, 26,* 444-450.

Baken, R.J. & Orlikoff, R.F. (2000). *Clinical measurement of speech and voice.* (second ed.) California: Singular Publishing Group Thomson Learning.

Booij, G. (1995). *The phonology of Dutch.* Oxford: Clarendon Press.

Bosshardt, H.G. (1999). Effects of Concurrent Mental Calculation on Stuttering, Inhalation and Speech Timing. *Journal of Fluency Disorders, 24,* 43-72.

Browman, C.P. & Goldstein, L. (1986). Towards an articulatory phonology. *Phonology Yearbook, 3,* 219-252.

Browman, C. P. & Goldstein, L. (1989). Articulatory gestures as phonological units. *Phonology, 6,* 201-251.

Browman, C.P. & Goldstein, L. (1990a). Tiers in articulatory phonology, with some implications for casual speech. In J.Kingston & M.E. Beckman (Eds.), *Papers in laboratory phonology: Between the grammar and the physics of speech* (pp. 341-376). Cambridge, UK: Cambridge University Press.

Browman, C.P. & Goldstein, L. (1990b). Gestural specification using dynamically-defined articulatory structures. *Journal of Phonetics, 18,* 299-320.

Browman, C.P. & Goldstein, L. (1990c). Representation and reality: Physical systems and phonological structure. *Journal of Phonetics, 18,* 411-424.

Browman, C.P. & Goldstein, L. (1992). Articulatory phonology: An overview. *Haskins Laboratories Status Report on Speech Research, SR-111,* 23-42.

Browman, C.P. & Goldstein, L. (1997). The gestural phonology model. In W.Hulstijn, H.F.M. Peters, & P.H.H.M. van Lieshout (Eds.), *Speech production: Motor control, brain research and fluency disorders* (pp. 57-71). Amsterdam, the Netherlands: Elsevier.

Byrd, D. (1996). Influences on articulatory timing in consonant sequences. *Journal of Phonetics, 24,* 209-244.

Byrd, D. & Tan, C.C. (1996). Saying consonant clusters quickly. *Journal of Phonetics, 24,* 263-282.

Caruso, A.J. (1988). Kinematic analysis of Multiple Movement Coordination during Speech in Stutterers. *Brain, 111,* 439-455.

Coover, J.E. (1923). A method of teaching typewriting based upon a psychological analysis of expert typing. *National Education Association, 61,* 561-567.

De Nil, L.F. (1995). The influence of phonetic context on temporal sequencing of upper lip, lower lip, and jaw peak velocity and movement onset during bilabial consonants in stuttering and nonstuttering adults. *Journal of Fluency Disorders, 20,* 127-144.

Diepstra, H.D., Huinck, W.J., Hulstijn, W., & Peters, H.F.M. (2001). The effect of word length on speech reaction time and word duration in the fluent word production of stuttering and nonstuttering subjects. In: B. Maassen, W. Hulstijn, R.D. Kent, H.F.M. Peters, & P.H.H.M. van Lieshout (Eds.), *Speech motor control in normal and disordered speech* (pp. 245-248). Nijmegen, the Netherlands: Vantilt.

Erikson, C.W., Pollack, M.D., & Montague, W.E. (1970). Implicit speech: Mechanism in perceptual encoding. *Journal of Experimental Psychology, 84,* 502-507.

Fowler, C.A. (1980). Coarticulation and theories of extrinsic timing. *Journal of Phonetics, 8,* 113-133.

Fowler, C.A. & Saltzman, E. (1993). Coordination and coarticulation in speech production. *Language and Speech, 36,* 171-195.

Fujimura, O. (1986). Relative invariance of articulatory movements: An iceberg model. In J.S.Perkell & D.H. Klatt (Eds.), *Invariance and and varibility in speech processes* (pp. 226-234). Hillsdale, NJ: Lawrence Erlbaum.

Fulk, R.D. (1999). Evaluating the evidence for lengthening before homorganic consonant clusters in the ormulum. In G.F.Car, W. Harbert, & L. Zhang (Eds.), *Interdigitations: Essays for Irmengard Rauch* (pp. 201-209). New York: Peter Lang.

Henry, F.M. (1980). Use of simple reaction time in motor programming studies: A reply to Klapp, Wyatt, and Lingo. *Journal of Motor Behavior, 12,* 163-168.

Hertrich, I., & Ackermann, H. (1995). Coarticulation in slow speech: Durational and spectral analysis. *Language and Speech, 38,*159-187.

Howell, P. & Vause, L. (1986). Acoustic analysis and perception of vowels in stuttered speech. *Journal of the Acoustic Society of America, 79,* 1571-1579.

Howell, P., Au-Yeung, J., & Sackin, S. (2000). Internal structure of content words leading to lifespan differences in phonological difficulty in stuttering. *Journal of Fluency Disorders, 25,* 1-20.

Huinck, W.J., Wouters, E.H.A., Hulstijn, W., & Peters, H.F.M. (2001). Differences between stuttering and non-stuttering people: Comparing diadochokinesis, sentences repetition and reaction time tasks. In: B. Maassen, W. Hulstijn, R.D. Kent, H.F.M. Peters, & P.H.H.M. Van Lieshout (Eds.), *Speech motor control in normal and disordered speech* (pp. 311-315). Nijmegen, the Netherlands: Vantilt.

Jakielski, K.J. (2002). A new method for measuring articulatory complexity. Paper presented at the 2002 *American Speech-Language-Hearing Association Annual Convention*, Atlanta, Georgia. 2002.

Jared, D. & Seidenberg, M.S. (1990). Naming multisyllabic words. *Journal of Experimental Psychology: Human Perception and Performance, 16,* 92-105.

Klapp, S.T., Anderson, W.G., & Berrian, R.W. (1973). Implicit speech in reading: Reconsidered. *Journal of Experimental Psychology, 100,* 368-374.

Klapp, S.T. (1995). Motor response programming during simple choice reaction time: The role of practice. *Journal of Experimental Psychology: Human Perception and Performance, 21,* 1015-1027.

Klapp, S.T. (1996). Reaction time analysis of central motor control. In H.N.Zelaznik (Ed.), *Advances in motor learning and control* (pp. 13-35). Champaign, IL: Human Kinetics.

Levelt, W.J.M. (1989). *Speaking: from Intention to Articulation.* Cambridge MA: MIT Press.

Levelt, W.J.M., Roelofs, A., & Meyer, A.S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences, 22,* 1-75.

Levelt, W.J.M. & Wheeldon, L. (1994). Do speakers have access to a mental syllabary? *Cognition, 50,* 239-269.

Liss, J. & Weismer, G. (1993). Qualitative acoustic analysis in the study of motor speech disorders. *Journal of Acoustical Society of America, 92,* 2984-2987.

Löfqvist, A. (1990). Speech as audible gestures. *Working papers in Logopedics and Phoniatrics, 6,* 58-98.

Löfqvist, A. & Yoshioka, H. (1981). Interarticulator programming in obstruent production. *Phonetica: International Journal of Speech Science, 38,* 21-34.

Munhall, K.G., Fowler, C.A., Hawkins, S., & Saltzman, E. (1992). "Compensatory shortening" in monosyllables of spoken English. *Journal of Phonetics, 20,* 225-239.

Munhall, K.G. & Löfqvist, A. (1992). Gestural aggregation in speech: Laryngeal gestures. *Journal of Phonetics, 20,* 111-126.

Nijland, L., Maassen, B., van der Meulen, S., Kraaimaat, F., & Schreuder, R. (in press). Planning of syllables by children with developmental apraxia of speech. *Clinical Linguistics and Phonetics.*

Öhman, S. (1966). Coarticulation in VCV utterances: Spectrographic measurements. *Journal of the Acoustical Society of America, 77,* 640-648.

Peters, H.F.M., Hulstijn, W., & Starkweather, C.W. (1989). Acoustic and physiological reaction times of stutterers and nonstutterers. *Journal of Speech and Hearing Research, 32,* 668-680.

Postma, A., Kolk, H. J., & Povel, D. J. (1990). Speech Planning and Execution in Stutterers. *Journal of Fluency Disorders, 15,* 49-59.

Riley, A. (1972). A stuttering severity instrument for children and adults. *Journal of Speech and Hearing Disorders, 37,* 314-322.

Robb, M.P. & Blomgren, M. (1997). Analysis of F2 transitions in the speech of stutterers and nonstutterers. *Journal of Fluency Disorders, 22,* 1-16.

Roelofs, A. (2002a). Syllable structure effects turn out to be word length effects: Comment on Santiago et al. (2000). *Language and Cognitive Processes, 17,* 1-13.

Roelofs, A. (2002b). Spoken language planning and the initiation of articulation. *The Quarterly Journal of Experimental Psychology, 55a,* 465-483.

Rosenbaum, D.A. (1991). *Human motor control.* San Diego: Academic Press, Inc.

Saltzman, E. (1986). Task dynamic coordination of the speech articulators: A preliminary model. In H. Heuer & C. Fromm (Eds.), *Generation and modulation of action patterns* (pp. 129-144). Berkin/Heidelberg: Springer-Verlag.

Saltzman, E. (1995). Intergestural timing in speech production: Data and modeling. In *Proceedings of the XIIIth international congress of phonetic sciences* (pp. 85-91).

Saltzman, E. & Kelso, J.A.S. (1987). Skilled actions: A task dynamic approach. *Psychological Review, 94,* 84-106.

Saltzman, E. & Munhall, K.G. (1989). A dynamical approach to gestural patterning in speech production. *Ecological Psychology, 1*, 333-382.

Santiago, J., MacKay, D.G., Palma, A., & Rho, C. (2000). Sequential activation processes in producing words and syllables: Evidence from picture naming. *Language and Cognitive Processes, 15,* 1-44.

Schiller, N.O. (1997). *The role of the syllable in speech production.* Dissertation, University of Nijmegen.

Schiller, N.O., Meyer, A.S., & Levelt, W.J.M. (1997). The syllabic structure of spoken words: Evidence from the syllabification of intervocalic consonants. *Language and Speech, 40,* 103-140.

Schmidt, R.A. & Lee, T.D. (1999). *Motor control and learning. A behavioral emphasis.* Champaign: Human Kinetics.

Shaiman, S. (2001). Kinematics of Compensatory Vowel Shortening: The Effect of Speaking Rate and Coda Composition on Intra- and Inter-Articulatory Timing. *Journal of Phonetics, 29,* 89-107.

SPSS Inc. (2000). SPSS for windows, release 10 [Computer software].

Sternberg, S., Monsell, S., Knoll, R.L., & Wright, C.E. (1978). The latency and duration of rapid movement sequences: Comparisons of speech and typewriting. In G.A. Stelmach (Ed.), *Information processing in motor control and learning* (pp. 117-152). New York: Academic Press.

Stoet, G. & Hommel, B. (1999). Action planning and the temporal binding of response codes. *Journal of Experimental Psychology-Human Perception and Performance*, 25, 1625-1640

Tiffany, W.R. (1980). The effect of syllable structure on diadochokinetic and reading rates. *Journal of Speech and Hearing Disorders, 23,* 894-908.

Tjaden, K. (1999). Can a model of overlapping gestures account for scanning speech patterns? *Journal of Speech, Language, and Hearing Research, 42*, 604-617.

Tjaden, K. (2000a). A preliminary study of factors influencing perception of articulatory rate in Parkinson disease. *Journal of Speech, Language, and Hearing Research, 43,* 997-1010.

Tjaden, K. (2000b). An acoustic study of coarticulation in dysarthric speakers with Parkinson disease. *Journal of Speech, Language, and Hearing Research, 43,* 1466-1480.

Tjaden, K. & Weismer, G. (1998). Speaking-rate-induced variability in F2 trajectories. *Journal of Speech, Language, and Hearing Research, 41,* 976-989.

Van der Merwe, A. (1997). A theoretical framework for the characterization of pathological speech sensorimotor control. In M.R.McNeil (Ed.), *Clinical management of sensorimotor speech disorders* (pp. 1-25). New York: Thieme Medical Publishers Inc.

Van Lieshout, P.H.H.M. (1995). *Motor planning and articulation in fluent speech of stutterers and nonstutterers.* PhD Dissertation, University of Nijmegen.

Van Lieshout, P.H.H.M., Hulstijn, W., & Peters, H.F.M. (1996a). From planning to articulation: What differentiates a person who stutters from a person who does not stutter. *Journal of Speech and Hearing Research, 39,* 546-564.

Van Lieshout, P.H.H.M., Hulstijn, W., & Peters, H.F.M. (1996b). Speech production in people who stutter: Testing the motor plan assembly hypothesis. *Journal of Speech and Hearing Research, 39,* 76-92.

Van Lieshout, P. & Moussa, W. (2000). The assessment of speech motor behaviors using electromagnetic articulography. *The Phonetician,* 81, 9-22.

Van Lieshout, P.H.H.M., Peters, H.F.M., & Bakker, A.J. (1997). En route to a speech motor test: A first halt. In W.Hulstijn, H.F.M. Peters, & P.H.H.M. Van Lieshout (Eds.), *Speech production: Motor control, brain research and fluency disorders* (pp. 463-471). Amsterdam: Elsevier Science BV.

Van Lieshout, P.H.H.M., Starkweather, C.W., Hulstijn, W., & Peters, H.F.M. (1995). Effects of linguistic correlates of stuttering on EMG activity in nonstuttering speakers. *Journal of Speech, Language, and Hearing Research, 38,* 360-372.

Weismer, G., Tjaden, K., & Kent, R.D. (1995). Speech production theory and articulatory behavior in motor speech disorders. In F. Bell-Berti & L.J. Raphael (Eds.), *Producing speech: Contemporary issues (For Katherine Safford Harris)* (pp. 35-50). New York: American Institute of Physics Press.

Wheeldon, L. & Levelt, W.J.M. (1995). Monitoring the time course of phonological encoding. *Journal of Memory and Language, 34*, 311-334.

Wingate, M.E. (1964). A standard definition of stuttering. *Journal of Speech and Hearing Disorders, 29,* 484-488.

Wingate, M.E. (1969). Stuttering as a phonetic transition defect. *Journal of Speech and Hearing Disorders, 34,* 107-108.

Yaruss, J.S. & Conture, E.G. (1993). F2 transitions during sound/syllable repetitions of children who stutter and predictions of stuttering chronicity. *Journal of Speech and Hearing Research, 36,* 883-896.

Young, E.C. & Thompson, C.K. (1987). An experimental analysis of treatment effects on consonant clusters and ambisyllabic consonants in two adults with developmental phonological problems. *Journal of Communication Disorders, 20,* 137-149.

Zebrowski, P., Conture, E.G., & Cudahy, E. (1985). Acoustic analysis of young stutterers' fluency: Preliminary observations. *Journal of Fluency Disorders, 10,* 173-192.

Ziegler, W. & Maassen, B. (in press). The role of the syllable in disorders of spoken language production. In B. Maassen, R.D. Kent, H.F.M. Peters, P.H.H.M. van Lieshout, & W. Hulstijn (Eds.), *Speech motor control in normal and disordered speech.* Oxford, UK: Oxford University Press.

Zsiga, E.C. (1994). Acoustic evidence for gestural overlap in consonant sequences. *Journal of Phonetics, 22,* 121-140.

Zsiga, E.C. (2000). Phonetic alignment constraints: Consonant overlap and palatalization in English and Russian. *Journal of Phonetics, 28,* 69-102.

Chapter Seven

Differences between stuttering and non-stuttering people: Comparing diadochokinesis, sentence repetition and reaction time tasks

Huinck, W.J., Wouters, E.H.A., Hulstijn, W., & Peters, H.F.M. (2001).
In B. Maassen, W. Hulstijn, R.D. Kent, H.F.M. Peters, & P.H.H.M. Van Lieshout (Eds.), Speech motor control in normal and disordered speech, pp. 311-315. Nijmegen, the Netherlands: Vantilt.

Abstract

To find dissimilarities between Person Who Stutter (PWS) and Persons Who do not Stutter (PWNS) we tested in 14 PWS and 14 PWNS the effect of word length, word meaning, sentence repetition, consonant cluster place and diadochokinesis. Acoustic reaction times and word durations were measured. Overall, PWS showed significantly longer reaction times and word durations than PWNS. Only in the reaction time data of the word meaning task the interaction between group and the main factor word versus non-word was significant. Furthermore, the largest percentage difference between the two groups was found in the Mean Absolute Deviation of the utterance duration in the sentence repetition task.

Introduction

This study is related to a research project in which the effect of therapy on speech production processes in stuttering adults will be assessed. Prior to the pre and post therapy comparison, it is necessary to find relevant tasks that differentiate between Persons Who Stutter (PWS) and Persons Who do Not Stutter (PWNS) and that show effects of manipulations on both groups so that possible changes that result from therapy can be detected. Differences between PWS and PWNS have been studied many times in different tasks. Some researchers, for instance, assessed the effect of word length on both groups (*e.g.* Soderberg 1966; Wingate, 1967; Peters, Hulstijn & Starkweather, 1989, van Lieshout, Hulstijn, Peters 1996a, 1996b). Others studied linguistic effects (*e.g.* van Lieshout, Starkweather, Hulstijn, Peters 1995), the effect of tongue twisters (*e.g.* Hulit & Loyd, 1987; Postma, Kolk, & Povel 1990; Wells & Moore 1990) or diadochokinesis. Many times though, researchers found contrary results as for example in Diadochokinesis tasks. Some of these studies showed significant differences between the two groups (Rickenberg, 1956; Devenny, 1990) whereas other studies did not find such differences (Chworowsky, 1951; Dworkin & Culatta, 1985). In word length studies, some researchers found that longer words elicit stutters more frequently (*e.g.* Soderberg 1966; Wingate, 1967) or that word length affects the reaction times of PWS more than PWNS (*e.g.* Peters, Hulstijn & Starkweather 1989) but others did not find such an effect (*e.g.* Van Lieshout, Hulstijn & Peters 1996). These studies are hard to compare with each other because different experimental designs were used, different groups were formed and different stimuli were selected. In the present study we tested the effect of word length, word meaning, sentence repetition of two different sentences, cluster place and diadochokinesis on 14 PWS and 14 PWNS. These tasks were organized in two types. Speech execution rate (diadochokinesis) and stability (repetition) were included to differentiate between the two groups. Effects of word meaning, word length and cluster type were included to focus on the interaction effects between group and the specific task level.

Methods

Subjects

Fourteen Persons who stutter (PWS) and fourteen persons who do not stutter (PWNS) participated in these experiments. The two groups were matched as closely as possible on age, gender and educational level. All subjects had normal hearing acuity, normal language and voice quality and were native speakers of Dutch. Mean age of PWS was 23.4 years (ranging between 17-36 year) and for PWNS 24.9 years (ranging between 17-36 year). All subjects gave their informed consent prior to participating in the experiment.

Tasks

Diadochokinetic

Diadochokinetic rates were obtained for each subject on the following sequences /pə/, /təkə/, and /pətəkə/. Each subject was asked to repeat as quickly as possible the sequences during respectively 5, 7 and 9 seconds.

Sentence Repetitions

Subjects had to repeat two different sentences. In sentence 1) *'Pieten pikken peperdure papavers'* (/pitən pIkən pe:pɔrdy:rə pɑpa:vɔrs/) each word started with the voiceless plosive /p/. In sentence 2) *'Biggen bieden bavianen bananen'* (/bIχən bi:dən ba:via:nən ba:na:nən/) each word began with the voiced plosive /b/. For the first vowel of each word we avoided the rounded vowels. Both sentences consisted of eleven syllables.

Word meaning

Sixty-four bi-syllabic stimuli were selected. Half of these stimuli had stress on the first syllable and the other half on the second syllable. Of both stress types half were low frequent Dutch words and the other half were non-words. The stress on the non-words was intuitively specified and it is therefore very well possible that some subjects put stress on the other syllable. However, we tried to select the stimuli as much as possible balanced for stress to correct for possible stress effects. Furthermore, each stimulus was presented in both a delayed and an immediate condition.

Word length

By varying word length systematically we investigated whether or not word length affects speech planning and production and if so, whether or not this effect is stronger for PWS than for PWNS. There were 54 different stimuli: 18 of these stimuli were mono-syllabic words, 18 stimuli were bi-syllabic words and 18 stimuli were three-syllable words. The onset was either ‘p’, ‘b’, ‘d’, ‘t’, ‘a’, ‘o’. Each stimulus was presented both in an immediate and a delayed condition.

Consonant cluster

Consonant cluster effects were studied on five sequence types: CVC, CCVC, CVC#CVC, CCVC#CVC. Of each type 16 stimuli were selected. The clusters in the first syllable consisted of either ‘br’, ‘pr’, ‘tr’ or ‘dr’. The medial cluster consisted of either ‘gt’ or ‘gl’. Each second syllable ended with ‘uk’(/ʌk/). Furthermore, each sequence was included four times: with /e:/, /ɛ/, /i:/, and /I/. Each trial was presented in a delayed reading condition.

Instrumentation and procedure
Each subject was tested individually in the presence of two experimenters. The stimuli were presented on a monitor in a delayed and/or in an immediate reading paradigm. The warning signal was given immediately followed by the presentation of the stimuli (delayed reading condition), or the stimuli were presented simultaneously with the go-signal (immediate reading condition). The subject had to respond after the go-signal at the instructed way (this was either as quickly as possible, in the reaction time tasks, or at a normal speaking rate). The end of the response time was indicated with a beep followed by a pause prior to the next trial.

Data analysis

Only utterances that were judged as fluently spoken were analyzed. Rt, wd and -in the diadochokinesis task- number of repetitions were calculated. Rt was measured as the interval between the go-signal and the first response of the acoustic signal. Wd was calculated as the time difference between the onset and the end of the acoustic signal and the number of repetitions was counted. The acoustic analysis program PRAAT was used for these analyses.

Statistical analysis

On each task a repeated measurement (SPSS, ANOVA GLM) was conducted with a varying number of within subjects factors (depending on the specific task) and group (PWS versus PWNS) as between subject factor. In the **Diadochokinesis task** the number of syllables in correct produced sequences was the dependent variable. There was one within subjects factor (*utterance type*: /pə/ /təkə/ and /pətəkə/). In the section **Sentence Repetition,** utterance duration and Mean Absolute Deviation (MAD)[22] of the utterance duration were analyzed. There was one within subject factor (*sentence type;* sentence 1 and 2). In the section **Word meaning** rt and wd were calculated. Three within subject factors (*word type*: word, non-word; *onset consonant* /b/, /d/, /p/, /t/; *reading condition*; delayed and immediate) were included. Rt and wd were also calculated in the **Word length** task. There were three within subject factors (*number of syllables*: 1, 2, 3; *onset consonant/vowel*: /b/, /d/, /o:/, /a:/, /t/, /p/; *reading condition*; delayed and immediate). For the analysis of the **Consonant Cluster task** rt and wd were determined. Two within subjects factors (*cluster type*: CVC, CCVC, CVCC, CVCCVC, CCVCCVC and *voicing*: voiced and unvoiced) were included. Furthermore, percentage difference between the two groups was calculated on each level of the tasks to get an impression of the value of tasks compared with each other.

---

[22] *The MAD is the absolute value of the deviations from the sample mean per subject.*

Results

Diadochokinesis task
PWS showed a smaller number of repetitions (mean=23.628 number of syllables per 5 seconds) than PWNS (mean=27.794 number of syllables per 5 seconds) on each sequence. This difference was significant ($F_{(1,26)}$=6.891, p=0.014). Although the difference between the two groups was not the same for each sequence, the interaction between sequence (/pə/ /təkə/ and /pətəkə/) and group was not significant ($F_{(2,25)}$=1.197, p=0.319).

Sentence Repetitions
We measured mean utterance duration and Mean Absolute Deviation (MAD) of the utterance duration. PWS had a longer utterance duration (mean utterance duration=221 ms) than PWNS (mean utterance duration=175 ms). This difference was significant ($F_{(1,25)}$=20.069, p=0.000). Furthermore, the MAD was calculated as a measure of variation. The MAD of the utterance duration for PWS (mean MAD=0.209) was larger than of PWNS (mean MAD=0.097). This difference was significant ($F_{(1,25)}$=25.378, p=0.000).

Word meaning task
Rt data. Words elicited shorter rt's (mean rt=583 ms) than non-words (mean rt=645 ms). This difference was significant ($F_{(1,25)}$=57.327, p=0.000). Furthermore, PWS were slower (mean rt = 675 ms) than PWNS (mean rt=553 ms). This difference was also significant ($F_{(1,25)}$=7.084, p=0.013). The difference between words and non-words was larger for PWS (difference=80 ms) than for PWNS (difference=42 ms) This interaction was significant ($F_{(1,25)}$=5.554, p=0.027).

Wd data. The wd of non-words (mean=493 ms) was longer than the word duration of words (mean=475 ms). This difference was significant ($F_{(1,25)}$=26.266, p=0.000).

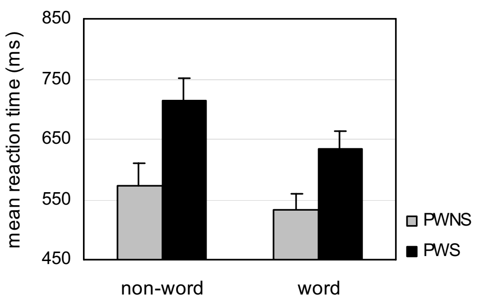


Figure 1. Mean reaction time (rt) and Standard Error for Persons who Stutter (PWS) and Persons who do not Stutter (PWNS) in word and non-word condition.

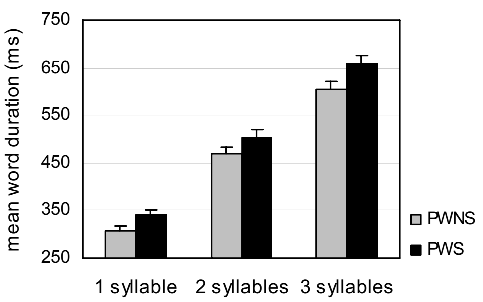


Figure 2. Mean word duration (wd) and Standard Error for Persons who Stutter (PWS) and Persons who do not Stutter (PWNS) in mono- bi- and tri-syllabic words.

Word length task
Rt data. As expected, rt increased with the number of syllables in the stimuli. The difference between mono-, bi-, and tri-syllabic words was significant ($F_{(2,23)}$=20.099, p=0.000). Furthermore, PWS were significantly ($F_{(1,24)}$=13.441, p=0.001) slower (mean rt=650 ms) than PWNS (mean rt=504 ms). The interaction between group and number of syllables was not significant ($F_{(2,23)}$=0.125, p=0.883).

Wd data. PWS showed longer word durations than PWNS (the mean wd for PWS=501 ms and for PWNS 460 ms). This difference was significant ($F_{(1,24)}$=4.49, p=0.045). The interaction between group and number of syllables approached significance ($F_{(2,23)}$=3.387, p=0.051). See also Figure 2.

*Table 1. F-values for the word duration (Wd) data and the reaction time (Rt) data (*p<0.05, **p<0.01).*

		F-value	
		Wd	*Rt*
Word length	Group	4.49 *	13.44 **
	Syllables	1481.34 **	20.10 **
	Group* Syllables	3.39	0.13
Consonant Cluster	Group	0.02	5.73 *
	Cluster	163.57 **	37.14 **
	Group*Cluster	0.54	1.35
Word meaning	Group	0.61	7.08 *
	Non-word	26.27 **	57.33 **
	Group*Non-word	0.17	5.55 *
Diadocho-kinesis	Group	-	6.89 *
	Sequence	-	17.52 **
	Group*Sequence	-	1.20
Repetition	Group	131.94 **	-
Repetition MAD	Group	35.15 **	-

Consonant Cluster task
Rt data. PWS showed longer reaction times than PWNS. The mean reaction time for PWS was 823 ms and for PWNS 769 ms. This difference was significant ($F_{(1,26)}$=5.729, p=0.024).As expected, the factor cluster type had effect on the reaction time. The different sequences elicited different reaction times in both PWS and PWNS. This difference was significant ($F_{(4,23)}$=37.142, p=0.000). The interaction between group and cluster type was not significant ($F_{(4,23)}$=1.353, p=0.281).

Wd data. As expected, the factor cluster type had a significant effect on word duration ($F_{(4,23)}$=163.572, p=0.00). However, although PWS showed longer word durations than PWNS, this difference was not significant ($F_{(1,26)}$=0.503, p=0.484). In Table 2, the percentages difference between the two groups and the mean results on each level of the tasks are presented. The MAD of the sentence repetition task shows the highest percentage difference (98.8 %) between the two groups. That is, the MAD of PWS is much larger than the MAD of PWNS. Furthermore, the diadochokinesis, word length and word meaning tasks elicit large differences between the two groups.

Table 2. Mean results of PWNS and PWS and the difference (Diff) and percentage difference (%Diff) calculated in relation to the mean of PWNS for each level of the diadochokinesis task (DDK), the word meaning task (WM), the word length task (WL) and the consonant cluster task (cluster).

Mean syllable duration (ms)		*PWNS*	*PWS*	*Diff*	*% diff*
DDK	/pə/	240	300	60	25.0
	/təkə/	201	240	39	14.4
	/pətəkə/	210	229	19	9.1
Mean Sentence duration (ms)					
Sentence Repetition		174	219	44	26.0
MAD		0.094	0.187	0.093	98.8
Mean reaction times (ms)					
Word meaning	Words	532	635	103	19.4
	Non-words	574	715	141	24.6
Word length	1 syllable	478	622	144	30.1
	2 syllables	497	639	142	28.6
	3 syllables	536	689	153	28.5
Cluster	CVC	793	849	56	7.1
	CCVC	738	755	17	2.3
	CVCC	808	810	2	0.3
	CVCCVC	645	722	77	11.9
	CCVCCVC	861	980	119	13.8
Mean word durations (ms)					
Word meaning	Words	466	483	17	3.5
	Non-words	486	500	14	2.9
Word length	1 syllable	308	340	32	10.4
	2 syllables	468	505	37	7.9
	3 syllables	605	659	54	8.9
Cluster	CVC	380	389	9	2.4
	CCVC	438	446	8	1.8
	CVCC	473	492	19	4.0
	CVCCVC	559	576	17	3.0
	CCVCCVC	622	656	34	5.5

Discussion

To compare PWNS with PWS on a set of tasks, we tested 14 PWS and 14 PWNS. Diadochokinesis and Sentence Repetition were mainly included to differentiate between the two groups and word meaning, word length and consonant cluster was varied to find possible interactions between the two groups. Diadochokinesis and Sentence Repetition showed significant effects of group. The groups showed a large percentage of difference on these tasks. Surprisingly, however, was the large percentage difference (98.8%) in the MAD. That is, PWS show more variation in their speech production. Furthermore, PWS were always slower than PWNS and their word duration was always longer. Except for the word duration data in the *Consonant cluster* task, each task showed significant differences between the two groups. The tested factor (number of syllables, word meaning, consonant cluster and sequence) was not always significant. The interaction between group and the tested factor was only significant in the reaction time data of the *word meaning* task.

Taken together, these results imply that the fluent speech of PWS shows a general difference from the speech of PWNS in word duration and reaction time. Extra loads as an increase in word length, and cluster variation affects both groups in the same way. Only one manipulation, the difference between words and non-words, was found to have a proportionally larger effect for PWS than for PWNS.

The tasks that show significant differences between the two groups and the task that shows a significant interaction between group and the particular main factor might be useful tasks for the pre-, post therapy comparison.

References

Chworowsky, C.R. (1951). A comparative study of the diadochokinetic rates of stutterers and non-stutterers in speech related and non-speech related movements. *Speech Monographs, 19,* 192.

Devenny, D.A., Silverman, W., Balgley, H., Wall, M.J., & Sidtis, J.J. (1990). Specific motor abilities associated with speech fluency in down's syndrome. *Journal of Mental Deficiency Research, 34,* 437-443.

Dworkin, J.P., & Culatta, R.A. (1985). Oral structural and neuromuscular characteristics in children with normal and disordered articulation. *Journal of Speech and Hearing Disorders, 50,* 150-156.

Hulit, L.M. (1987). Influences of suggestion on nonflluencies of normal speakers: a pilot study. *Perceptual and Motor Skills, 65,* 164-164.

Peters, H.F.M., Hulstijn, W., & Starkweather, W.C. (1989). Acoustic and physiological reaction times of stutters and nonstutters. *Journal of Speech and Hearing Research,* 668-680.

Peters, H.F.M., Hulstijn, W., & Van Lieshout, P.H.H.M. (1985). Towards a nijmegen speech motor test. In: C.W. Starkweather, & H.F.M. Peters (Eds.), (pp. 15-19). Nijmegen: Nijmegen University Press.

Postma, A., Kolk H.J., & Povel D.J. (1990). Speech planning and execution in stutterers. *Journal of Fluency Disorders, 15,* 49-59.

Rickenberg, H.E. (1956). Diadochokinesis in stutters and non-stutters. *Journal of the Medical Society of New Yersey, 53,* 324-326.

Soderberg, G.A. (1966). The relations of stuttering to word length and word frequency. *Journal of Speech and Hearing Research, 9,* 589.

Tabachnick, B.G., & Fidell L.S. (1996). Using multivariate statistics. New York: HarperCollins College Publishers.

Van Lieshout, P.H.H.M., Hulstijn, W., & Peters, H.F.M. (1996a). Speech production in people who stutter: testing the motor plan assembly hypothesis. *Journal of Speech and Hearing Research,* 76-92.

Van Lieshout, P.H.H.M., Hulstijn, W., & Peters, H.F.M. (1996b). From planning to articulation: what differentiates a person who stutters from a person who does not stutter. *Journal of Speech and Hearing Research,* 546-564.

Van Lieshout, P.H.H.M., Starkweather, C.W., & Hulstijn, W. (2000). effects of linguistic correlates of stuttering on EMG activity in nonstuttering speakers. *Journal of Speech and Hearing Research, 38,* 360-372.

Wells, B.G., & Moore W.H. (1990). EEG alpha asymmetries in stutterers and non-stutterers: Effects of Linguistic vaiables on hemispheric processing and fluency. *Neuropsychologia, 28,* 1295-1305.

Wingate, M.E. (1967). Stuttering and word length. *Journal of Speech and hearing Reasearch, 10,* 146-152.

Chapter Eight

General discussion

General Discussion

There are many approaches to the treatment of stuttering, but the lack of empirically motivated research into therapy outcomes prompted the longitudinal and elaborated evaluation of three accepted treatment approaches (Chapter 2). This type of research has gained importance with the increasing emphasis on evidence-based practice (EBP). An additional advantage of systematic evaluations is that the objective assessment of specific stuttering treatments also elucidates factors that contribute to desired outcomes (Blomgren, Roy, Callister & Merrill, 2005). Thus, by studying treatment efficacy concurrently with specific skills in both persons who stutter (PWS) and those that do not (PWNS), the underlying mechanisms of the speech deficit can be revealed.

In stuttering outcome studies, it is not possible to randomly assign the participants to different types of stuttering therapy since motivation is the key to treatment success. At the same time, this hampers direct comparisons of different therapy programmes. Because in Chapter 2 we presented the results of various interventions, we were tempted to draw conclusions about which of the evaluated therapies was best. However, because of -among other restrictive factors- the lack of randomization instead, questions about why some clients were very successful and others less so or even failed to benefit, led us to formulate new research questions. Our interest was twofold. We wished to (1) improve our understanding of what type of 'stuttering profile' profits most from what type of stuttering therapy, and (2) increase our insights into the underlying (timing) mechanisms of stuttering.

The need for the classification of stuttering profiles originated from the well-known heterogeneity in the stuttering population (*e.g.* Clutter & Freeman, 1984; Schwartz & Conture, 1988), which made the efficacy of stuttering therapies unpredictable. After having found positive effects for the evaluated stuttering programmes (Chapter 2), we were indeed able to demonstrate that individuals with different stuttering profiles responded differently to a selected treatment (Chapter 3). The results illustrated that different treatment approaches are indicated for different profiles. In addition, if clients with different stuttering profiles are assigned to specific programmes, the treatment's efficacy will be enhanced.

As this was just a first step into this field, we felt that our subsequent research efforts should be aimed at a further unravelling of the different stuttering profiles and at the relationship between these profiles and specific types of treatment approaches (*e.g.* individual versus group therapy and fluency-enhancing versus cognitive behavioural therapy). This idea was strengthened by the recent work of Stager et al. (2005) who investigated the effect of treatment with medications affecting dopaminergic and serotonergic mechanisms on fluency and anxiety in PWS. Their results indicated that it could be dopaminergic mechanisms that underlie the pathophysiology of stuttering. Stager and co-workers also suggested that the stuttering population might be subdivided, based on individual responses to dopamine antagonist medication.

It had become apparent that the choice of a specific type of therapy programme should be based on the specific profile of the help seeker. Then, if a client could be classified founded on a set of specified indications, this would enhance the objectivity of the diagnostic procedure, which, in turn, would probably also enhance the overall efficacy of the programme. This implies that existing therapies should be fine-tuned to fit the client's stuttering profile, which is no easy task. With our evaluation of Chapter 3 we, in line with Miller and Watson (1992), demonstrated that there is no relationship between stuttering severity and the severity of negative emotions and cognitions associated with dysfluency, which means that PWS with severe impediments are not necessarily emotionally affected. Although we did not explicitly investigate how different types of stuttering profiles responded to the chosen treatment, the results presented in Chapter 3 indicated that interventions with a structured follow-up programme are most successful in PWS with high scores on the stuttering severity index. The largest gain as well as the largest regression were found in the PWS with the severer levels of stuttering, suggesting that, after therapy completion, severe stutterers may need additional attention and coaching in the follow-up period. This, in contrast, does not seem to apply to PWS whose dysfluency is associated with severe emotional/attitudinal problems. Here, the findings suggested that it is likely that this type of speaker is capable of improving his/her attitude toward stuttering by forming more realistic beliefs about their speech.

Is therapy A better than therapy B? Clearly, clinicians, researchers and many of those who stutter are greatly interested in the efficacy of treatment programmes. Yet, the question is still hard to answer. What exactly *efficacy* or *effectiveness* in stuttering therapy is, has been elaborately discussed in the literature (see e.g. Bernstein Ratner, 2005; Bernstein Ratner & Healey, 1999; Bothe, 2003; Conture, 1996, 2001; Cordes, 1998; Craig, 2002; Curlee, 1993; Eichstaedt, Watt & Girson, 1998; Finn, 2003ab; Huinck & Donders 2005; Ingham, 1993, 2003; Ingam & Cordes, 1997; Langevin & Kully 2003; Onslow, 2003; Thomas & Howell, 2001; Yaruss, 1998ab; Yaruss, 2001; Yaruss & Quesal, 2004). Yet, defining the term in our context proved not to be as simple as in many medical treatments, where, for example, the effect of medication can be measured in blood samples or bio-parameters. Also, and as mentioned above, in efficacy research of stuttering therapy random assignment of participants is deemed unethical and unwise on account of the motivation/compliance issue. Moreover, appropriate control groups are difficult to find. Most adult PWS have already participated in one or more treatment programmes before, which makes them unsuitable as controls. These factors all prevented us from making sound comparisons between the three treatment programmes we evaluated in Chapter 2. Nevertheless, all three interventions proved to be effective. When, two years after treatment completion, all clients were asked whether they still benefited from the treatment, all but one respondents claimed they still did. Some relied more on speech techniques, others had become more confident. Thus, although we were unable to assess whether the various approaches had affected therapy outcomes differently, it did become evident that different approaches elicited different effects. However, Bernstein-Ratner (2005) argued that in psychotherapy specific techniques have been found to account for no more than 15% of the variance in therapy outcomes (also see Chwalisz, 2001; Lambert & Barley, 2001). On account of this minor percentage one may alternatively argue that PWS are perfectly able to choose the treatment that fits them best or to select the most appropriate and therefore most effective therapist (as has been suggested by Ahn and Wampold, 2001). If so, providing them with information about the existing types of stuttering programmes may suffice. It would be interesting if future research could corroborate this notion.

It is evident that many factors may play a role in the success of stuttering therapies (see e.g. Bothe, 2003; Finn, 2003; Ingham, 2003; Onslow, 2003; Thomas & Howell, 2001; Yaruss, 2001). However, one aspect that has not received much attention in the literature is how therapy programmes differ within different cultures. In the Netherlands, apart from programmes that have been designed in the Netherlands, therapists also use interventions that were originally developed in other -basically western- countries. Chapter 4 demonstrated that the Canadian Comprehensive Stuttering Program was similarly effective in a Dutch setting. Although rather small, the differences between ours and the Canadian culture in perceptions of self as a stutterer at F2 and the change in self-perceptions in our group are intriguing. Especially because clinicians from The Institute for Stuttering Treatment and Research (ISTAR) in Canada came to the Netherlands for this clinic, which implied that the differences in treatment outcome could not have been caused by differences in the attitudes and ideas of the attending clinicians. If future studies confirm our findings, these differences between the cultures, be they minor, should be taken into account when new therapy approaches from abroad are introduced in this country.

The evaluation of three treatment programmes, as described in Chapter 2, included a large set of tests (*e.g.* percentage stuttering, speech rate, self-evaluation scores, listeners' judgements on the speech of PWS and measurements of speech motor control). This large set of data allowed us to evaluate how the various efficacy instruments that were used, were related to each other. The study did raise the question whether there would be an easier, less time consuming way to study the efficacy of stuttering programmes using a smaller set of instruments. In Chapter 5 we found evidence for the validity of a School Mark (SM) assessment: its scores followed the same pattern of results as the other self-evaluation measures we had applied. The SM assessment showed significant correlations on the difference scores between pre- and post assessments with the percentage stuttered syllables, speech rate, and the judgments of listeners who scored the level of dysfluency in speech samples. On the other hand, SM assessment did not correlate with self-evaluation questionnaires and naturalness judgements and listeners' comfort scores. The SM thus proved a simple and effective alternative for the widely used standard instruments facilitating a fast and cost-effective treatment efficacy evaluation. The SM not only yielded a valid score, it also included the most critical goal of stuttering treatment: self-reported acceptability (Ingham and Cordes, 1997). It can be used as a measure of therapy progress within therapy sessions and also as a measure of therapy effectiveness directly after the end of the intervention.

Because of its succinctness the SM may not be the most sophisticated of instruments. It combines many subtle characteristics and details regarding the complex stuttering problem and this information will not be revealed when using this type of efficacy measure only. Some PWS, for instance, may have high expectations of the therapy to which they have committed themselves. When these expectations fail to materialize, feelings of disappointment might strongly decrease the SM score. This effect will vary for different (types of) stutterers and at different points in time.

Part II of this thesis focused on aspects of speech motor control in PWS and PWNS. Subscales of 'The Nijmegen Speech Motor Test' (NMST) were analysed to assess the *underlying* aspects of stuttering by looking at speech reaction time (RT) and word duration (WD). Many researchers have attempted to delineate the mechanisms that underlie stuttering, but to date no definitive suggestions have been put forward although many interesting potential sources have been proposed. Among these potential causes are learned behaviour (*e.g.* Bloodstein, 1969, 1995; Brutten & Shoemaker, 1967; Johnson, 1959, Shames & Sherrick, 1963) and disrupted processes in speech motor control (*e.g.* Adams, Freeman & Conture, 2003; Kent, 1984; Max, 2004; Van Lieshout, Peters, Starkweather & Hulstijn, 1993; Van Lieshout, Hulstijn & Peters, 2004; Wood, 1996; Zimmerman, 1980) and also a combination of the former two (*e.g.* Van Riper, 1982; Peters & Guitar, 1991). The lack of conclusive evidence in support of speech motor disruption (see e.g. Ingham, 1998) resulted in a shift in attention towards a more 'linguistic' vision on stuttering in which stuttering was assumed to be related to a linguistic planning impairment (Howell & Au-Yeung, 2002; Kolk, 1991; Kolk & Postma, 1997; Conture, 2001; Conture, Zackheim, Anderson, Pellowski, 2004). In one specific view, the Covert Repair Hypothesis, the impairment was related to mechanisms that are involved in the monitoring of these planning processes (see Levelt, 1989; Postma 2000). Recently, causes like overactivity of the basal ganglia (*e.g.* De Nil, Kroll, Kapur & Houle, 2000; Stager et al., 2005; Whu et al., 1997) have been tested. In contrast to the advocates of 'single-sided' views, there are those who favour a multi-factor approach, emphasising that various aspects can play a role in stuttering depending on a critical balance between existing capacities and communicative demands (Adams, 1990; Smith & Goffman, 2004; Starkweather & Gotwald, 1990). They also stress that speech motor and linguistic processes are much closely related and involved with each other than traditional speech production models suggest (Smith & Goffman, 2004; see also Van Lieshout, 1997).

In this thesis, we measured RTs and WDs to study how speech motor processes in PWS and PWNS could be challenged without directly addressing theories about the cause of stuttering. The two groups were tested with the Nijmegen Speech Motor Test (NSMT). The results revealed information about which speech tasks were most difficult to perform and which tasks best differentiated between the two groups (see Chapters 6 and 7). Complexity of the specific speech task proved to affect the speech production of PWS. Overall, the PWS were slower in their speech production, more variable in their speech output and as a group more heterogeneous than the fluent speakers. Moreover, when the dysfluent group was confronted with a more complex task, their RTs increased additionally and more dysfluencies and speech errors occurred, both relative to the simpler task and to the PWNS. Task complexity was manipulated by (1) increasing the complexity of coarticulation patterns (see Chapter 6) and by (2) among others, the lack of meaning of a word (nonwords) (Chapter 7). Of course, complexity of speech is a daring subject within the field of phonetics, linguistics and speech and language pathology. It is almost impossible to be complete when specifying the complexity of an utterance. The planning and production of a specific utterance is influenced by many facets of complexity, among which are place of articulation, manner of articulation, coarticulation and assimilation processes, and word and syllable frequencies. Hence, one should be very careful when drawing conclusions. Nevertheless, the interaction processes between stuttering and speech complexity are very interesting. Insights into these interactions might help us expand our understanding of pathological speech.

In Chapter 6 we describe that homorganic clusters (*i.e.* consonants in a cluster with the same place of articulation) across a syllable boundary gave rise to longer delays in RTs for the PWS compared to those of the PWNS, which was not the case for the heterorganic clusters (*i.e.* consonants in a cluster with different places of articulation). Based on concepts from Gestural Phonology (Goldstein & Fowler, 2003) we argued that homorganic clusters put greater demands (*i.e.* restrictions) on gestural coordination, which is especially critical if the last segment in the cluster forms the onset of a syllable/word (as was the case). Such extra demands may require a longer preparation time. Howell, Au-Yeung and Sackin (2000) found that consonant clusters at word-initial position increased the chance of stuttering whereas clusters at other positions did not have a significant influence on stuttering frequency. The authors related their finding to the stage of phonological developmental difficulty. They assumed that late developing sounds (*e.g.* consonant strings versus singletons) are not only relevant to child phonological development, but may also increase demands (i.e. delay processing time) during the pre-motor planning stages of these sounds in speakers of all ages. How the pronunciation complexity of clusters have impact on planning is not clear, nor if this is a phonetic rather than a phonological effect. This aspect will need to be addressed further in future research.

Chapter 7 showed that, in addition to the generally slower RTs, nonwords elicited an extra delay in the PWS group when compared to words with similar complexity. The role of language processes in speech production (and perception) has been discussed at length in the literature (see e.g. Seidenberg and McClelland, 1989, and Andrews and Scarratt, 1998) but the views on their relevance to explain the origin and maintenance of stuttering are far from consistent (see Bernstein-Ratner, 1997, and Conture, Zackheim, Anderson and Pellowski 2004, for a review). A set of studies by Rastatter and Dell (1985, 1987a, 1987b) showed that, among other differences, their PWS were significantly slower than the PWNS in a lexical decision task but not in a simple vocal reaction-time task. Bosshardt and Fransen (1996) tested online sentence processing in a self-paced word-by-word reading experiment in PWS and matched PWNS. Both groups showed a similar speed for word identification but PWS were found to be slower in making decisions about word categories, which the authors interpreted as indicative of delays in retrieving semantic information. In contrast, Van Lieshout, Hulstijn and Peters (1996) found no significant differences between PWS and matched controls in word- or symbol-naming tasks, suggesting that word-related planning processes did not seem to be a major issue for PWS. The only group difference the authors did observe was related to upper- and lower-lip muscle-activation time, which they attributed to delays in motor initiation as part of a compensatory motor-control strategy used by PWS to remain fluent. Packman, Onslow, Coombes and Goodwin (2001) also did not find problems in lexical processing in their PWS, which would be inconsistent with the idea that lexical retrieval problems are fundamental in stuttering as suggested, among others, by Postma and Kolk (1993) and Au Yeung and Howell (2002). The latter authors also provided several arguments against the interpretation of Packman and colleagues (2001).

Although there is no consistent idea about the role of lexical retrieval in stuttering, one way to explore its involvement is to contrast verbal items with or without a lexical representation. As Chapter 7 demonstrated, the use of words and nonwords yields a contrast. Nonwords have no meaning and therefore can be assumed to have no lexical entry point (but see Lukatela, Carello, Savic, Urosevic and Turvey, 1998, who showed how nonwords can provide access to the mental lexicon). It also means that in terms of phonological encoding there is no stored information about their segment and metrical structure. In normal speakers, this means that in verbal naming tasks (*i.e.* reading aloud) nonwords have to go through a grapheme-to-phoneme spelling process for the phonological string to be assembled (see Coltheart, Curtis, Atkins & Haller 1993), whereas existing, meaningful words follow a more direct and faster lexical route (see also Levelt and Wheeldon, 1994). Several functional neuroimaging studies also confirmed differences between words and nonwords (for a review, see Wilson, Leuthold Lewis, Georgopoulos and Pardo, 2005). Wilson et al. (2005) concluded that a neural distinction in word/nonword processing should be conceptualized as a dynamic difference in processing *time*. Nonwords showed an average delay of 100 ms in activation of the involved brain areas when compared to words, which could tentatively correspond to the assumed extra processing time required to assemble a phonological/phonetic plan for a nonword. Apart from differences in lexical retrieval and phonological encoding, a recent study also suggested that nonwords may induce a different type of articulation compared to similar sounds produced in real-word contexts (McClean, Tasko, & Runyan, 2004; Tasko & McClean 2004).

If we assume that lexical retrieval is delayed in PWS, nonwords could be assumed not to pose any problems as there is no lexical processing required. On the other hand, if the problem in PWS is in the actual assembly of a phonological/phonetic plan, the extra time required to complete this process for nonwords might have a larger impact on impeded speakers. If we take the RT differences between words and nonwords for normal speakers as the norm, we can make the following assumptions:

a. If in PWS lexical retrieval is problematic but phonological/phonetic encoding is not, PWS should show a larger RT for words but not for nonwords. Thus, the RTs in response to nonwords should actually be shorter for PWS than for PWNS.
b. If lexical retrieval is normal but phonological/phonetic encoding is problematic, PWS should show a larger increase in RT for nonwords in particular. This means that for PWS the difference in RTs between nonwords and words should be larger than those for PWNS.
c. If both processes are affected in PWS, RTs should be larger for both words and nonwords, rendering the actual difference between the two very similar to that of PWNS.

If nonword-word contrasts also affect speech motor preparation and execution processes, the exact nature of the response time changes is hard to predict, depending on individual nonword characteristics (Kendall et al., 2005). However, based on the work by Kendall's group and other researcher teams, RT and WD will be negatively correlated (shorter RT, longer WD and vice versa). This would not be obvious from effects that arise in earlier stages of word production as mentioned above.

Chapter 7 revealed that relative to words, nonwords elicited longer speech RTs and longer WDs. Moreover, when we compared in an additional analysis the RTs of all immediate reading tasks described in Part II, we found that the RTs in the nonword condition (all CVC sequences) exceeded the meaningful words. They were even longer than the bi-syllable and tri-syllable (meaningful) words in the word-length task. They were also longer than the homorganic and heterorganic nonwords in Chapter 6. The latter set of stimuli all started with a similar onset, /ba:/, which made them probably easier to prepare than the nonwords from Chapter 7. This not only applied for the PWS group; a similar effect of nonwords was found in the PWNS group. This finding corroborated our second (b) prediction: in PWS phonological/phonetic encoding is problematic but lexical retrieval is not.

But how can this be explained? When reading words, PWS take advantage of the lexical information that they stored in a buffer. The fact that this process did not elicit longer RTs suggests that there is no lexicalization problem in stuttering. Nonwords, on the other hand, force the speaker to go through a grapheme-to-phoneme spelling whereas words can be retrieved immediately from the lexicon (which should reduce the online computational load). If there were lexical issues in stuttering, these would have elicited a larger difference between the groups in the word condition. We found the opposite to be true with the difference between the two groups being larger in the nonword condition. This suggests that stuttering could be related to either phonological encoding (spelling out the word's metrical and segmental properties) or phonetic encoding (speech motor planning) or to both. Similar types of problems are implicated in apraxia of speech and researchers in this type of speech pathology are engaged in comparable discussions about the levels of information processing that are involved. According to Varley and Whiteside (2001) there is a direct route for frequent words between phonetic and lexical access or storage of whole words, although Ziegler (2001) argued against this idea and postulated that, rather than a lexical disorder, apraxia of speech is located on a much lower level.

That we did not find added delays in the PWS group in the cluster-complexity task can be explained by the fact that this task was only tested in the delayed-reading condition. Participants had more time to prepare the difficult clusters. On the other hand, there were no effects of cluster complexity when clusters were placed in coda position (Chapter 6). The effects occurred in the (articulatory less complex?) syllable-boundary condition. We are not convinced that these results reflect a real higher-level problem. Rather, the effect may just as well have something to do with lower-level processes. According to Ziegler (2003), for instance, speech motor control is task-specific. Evidence for this lies in clinical and experimental dissociations between speech and non-speech tasks in dysarthria and apraxia of speech. Ziegler argues that the production of nonwords is not really speech but rather a special type of oral motor task. If so, this implies that the problem PWS have with nonwords is in being limited in learning new oral motor tasks and that they therefore show more errors and longer

RTs. Why we found this effect in the immediate-reading condition only is unclear. Possibly, in the delayed-reading condition participants had more preparation time for phonological/phonetic encoding.

Another potential explanation for the delayed RTs in stuttering is 'strategy'. To maintain their fluency, PWS often apply specific strategies like reducing their speech rate and taking more time to produce speech. It could be argued that the differences between PWS and PWNS results from such tactical effects. However, there are two arguments against this interpretation: (1) this would have yielded delays in WDs as well but we only seldom found such delays, and (2) it does not explain the interaction between word type (word versus nonwords) and group.

How can the results of part I of this thesis be related to the results of part II? In part I the participants were divided into subtypes based on stuttering severity. In part II speech reaction times (RTs) proved to be a valid instrument to assess nonword-processing complexity. The next question was whether stuttering severity would be related to the RTs in the nonword condition, in other words, whether the longer RTs in relation to nonwords are symptoms of stuttering. If so, reaction time might also be an indication of stuttering severity. Our additional analysis to calculate the correlation between the percentage stuttered syllables and the RT difference between words and nonwords of 46 randomly selected PWS did not generate a significant association. However, we did find a correlation between speech rate (syllable per minute) and the percentage of stuttered syllables. The lower the speech rate, the higher the percentage of stuttered syllables. This correlation may indicate that PWS who produce many dysfluencies during speech automatically try to reduce these by controlling and slowing down their speech production.

Furthermore, Part II showed that stuttering involves problems at the level of phonological/phonetic (nonword delays) planning, but also at the execution phase (gestural overlap complexity). Although both findings seem to address different aspects of speech, we do think that they both support the view that stuttering is a motor problem rather than a linguistic one. In Part II both verbal naming tasks required more articulatory skills in the complex conditions. The homorganic conconant clusters and the nonwords were unknown and unautomatized combinations. The speech-production system had to rely on speech motor control and auditory, proprioceptive and sensory feedback during performance of the tasks. This combination of processes seemed to be problematic for the PWS population. We argue that when speech becomes more difficult or complex, extra mental resources are needed to remain fluent. In fluent speakers, speech production (including phonation, articulation, respiration and timing) is an unconscious and (to some degree) automatic process whereas in PWS it is not, resulting in dysfluencies. In stuttering therapy PWS learn how to control their speech and to become aware of the different processes that are involved in speaking. Speech techniques like reducing the rate of speech (with or without regulation of breathing) and smooth contact have shown to be very effective. The use of these techniques takes more time, which results in a more controlled use of the mental capacities or resources allowing fluency to be maintained. As soon as PWS try to convert this into an automatic process, they tend to relapse into dysfluent speech. This explains why PWS often relapse after therapy completion, especially in the long-term (one year). This also explains why we found more dysfluencies in the unstructured tasks like the monologue and the interview and fewer in the speech tasks included in the Nijmegen Speech Motor Test, which were more structured in nature.

To conclude, in this thesis we have attempted to bridge the gap between theories of stuttering and the practice of therapeutic interventions. Our theory showed that people who stutter lack the mental recourses to remain fluent in complex speech tasks, and, as a result, need more time and mental capacities to remain fluent. Speech techniques acquired in stuttering therapy seem to improve these timing aspects of stuttering. However, efficacy is related to different types of stuttering profiles and therapy programmes and diagnostic procedures should be adapted to fit these profiles.

References

Adams, M.R., Freeman, F.J., & Conture, E.G. (2003). Laryngeal dynamics of stutterers. In R.F. Curlee & W.H. Perkins (Eds.), *Nature and treatment of stuttering: New directions* (pp. 89-129). College-Hill Press.

Adams, M.R. (1990). The demands and capacity model I: Theoretical elaborations. *Journal of Fluency Disorders, 15,* 135-141.

Ahn, H., & Wampold, B.E. (2001). Where oh where are the specific ingredients? A meta-analysis of component studies in counseling and psychotherapy. *Journal of Counseling Psychology, 48,* 251-257.

Andrews, S., & Scarratt, D.R. (1998). Rule and analogy mechanisms in reading nonwords: Hough dou peapel rede gnew wirds. Journal of Experimental Psychology: Human Perception & Performance, 24, 1052–1088.

Au-Yeung, J., & Howell, P. (2002). Non-word reading, lexical retrieval and stuttering: Comments on Packman, Onslow, Coombes and Goodwin (2001). *Clinical Linguistics and Phonetics, 16,* 287-293.

Bernstein Ratner, N. (1997). Stuttering: a psycholinguistic perspective. In R.F. Curlee & G.M. Siegel (Eds.), *Nature and treatment of stuttering. New directions* (second ed., pp. 99-127). Boston: Allyn and Bacon.

Bernstein Ratner, N. & Healey, E.C. (1999). Bridging the gap between stuttering research and practice: An overview. In N. Bernstein Ratner & E.C. Healey (Eds.), *Stuttering research and practice: Bridging the gap* (pp. 1-12). Mahwah, New Jersey: Lawrence, New Jersey.

Bernstein Ratner, N. (2005). Evidence-based practice in stuttering: Some questions to consider. *Journal of Fluency Disorders, 30,* 163-188.

Blomgren, M., Roy, N., Callister, T., & Merrill, R.M. (2005). Intensive stuttering modification therapy. A multidimensional assessment of treatment outcomes. *Journal of Speech, Language, and Hearing Research, 48,* 509-523.

Bloodstein, O. (1969). *Handbook on Stuttering*. Chicago: National Easter Seal Society for Crippled Children and Adults.

Bloodstein, O. (1995). *A handbook on stuttering, Fifth edition*. San Diego: Singular Publishing Group, Inc.

Bosshardt, H.G., & Fransen, H. (1996). Online sentence processing in adults who stutter and adults who do not stutter. *Journal of Speech and Hearing Research, 39,* 785-797.

Bothe, A.K. (2003). Evidence-based treatment of stuttering: V. The art of clinical practice and the future of clinical research. *Journal of Fluency Disorders, 28,* 247-258.

Brutten, E.J., & Shoemaker, D.J. (1967). *The Modification of Stuttering*. Englewood Cliffs NJ: Prentice Hall.

Chwalisz, K. (2001). A common factors revolution: Let's not "Cut off our discipline's nose to spite its face". *Journal of Consulting Psychology, 48,* 262-267.

Clutter, M.H., & Freeman, F.J. (1984). Stuttering: The six blind men revisited. *Journal of Fluency Disorders, 9,* 89-92.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. (2nd ed.) New Jersey: Lawrence Erlbaum Associates.

Coltheart, M., Curtis, B., Atkins, P., & Haller, M. (1993). Models of reading aloud: Duak-route and parallel-distributed-processing approaches. *Psychological Review, 100,* 589-608.

Conture, E.G. (1996). Treatment efficacy: Stuttering. *Journal of Speech and Hearing Research, 39,* 518-526.

Conture, E.G. (2001). *Stuttering: Its nature, diagnosis, and treatment.* Boston: Allyn & Bacon.

Conture, E.G., Zackheim, C.T., Anderson, J.D., & Pellowski, M.W. (2004). Linguistic processes and childhood stuttering: many's a slip between intention and lip. In B. Maassen, R.D. Kent, H.F.M. Peters, P.H.H.M. Van Lieshout, & W. Hulstijn (Eds.), *Speech motor control in normal and disordered speech* (pp. 253-281). Oxford: University Press.

Cordes, A.K., & Ingham, R.J. (1998). *Treatment efficacy for stuttering.* San Diego: Singular Publishing Group, Inc.

Craig, A. (2002). Fluency outcomes following treatment for those who stutter. *Perceptual and Motor Skills, 94,* 772-774.

Curlee, R.F. (1993). Evaluating treatment efficacy for adults: Assessment of stuttering disability. *Journal of Fluency Disorders, 18,* 151-158.

De Nil, L.F., Kroll, R.M., Kapur, S., & Houle, S. (2000). A positron emission tomography study of silent and oral single word reading in stuttering and nonstuttering adults. *Journal of Speech, Language, and Hearing Research*, *43*, 1038–1053.

Eichstaedt, A., Watt, N., & Girson, J. (1998). Evaluation of the efficacy of a stutter modification program with particular reference to two new measures of secondary behaviors and control of stuttering. *Journal of Fluency Disorders, 23,* 231-246.

Finn, P. (2003). Evidence-based treatment of stuttering: II. Clinical significance of behavioral stuttering treatments. *Journal of Fluency Disorders, 28,* 209-218.

Finn, P. (2003). Addressing generalization and maintenance of stuttering treatment in the schools: a critical look. *Journal of Communication Disorders, 36,* 153-164.

Goldstein, L., & Fowler, C.A. (2003). Articulatory Phonology: A phonology for public language use. Internet: http://seneca.uab.es/pilarprieto/aula/Articulatory-phonology.pdf.

Howell, P., & Au-Yeung, J. (2002). The EXPLAN theory of fluency control applied to the diagnosis of stuttering. In Fava (Ed.), *Pathology and therapy of speech disorders* (Amsterdam: John Benjamins.

Howell, P., Au-Yeung, J., & Sackin, S. (2000). Internal structure of content words leading to lifespan differences in phonological difficulty in stuttering. *Journal of Fluency Disorders, 25,* 1-20.

Huinck, W.J., & Donders, A.R.T. (2005). Meten van effectiviteit van stottertherapie. Peters, H.F.M., Bastiaanse, R., Van Borsel, J., Dejonckere, Ph., Jansonius-Schultheiss, K., Van der Meulen, S., and Mondelaers, B.J.E. [B7.3.3], 1-18. Houten, Bohn Stafleu Van Loghum. Handboek Stem-, Spraak-, en Taalpathologie.

Ingham, J.C., & Cordes, A.K. (1997). Self-measurement and evaluating stuttering treatment efficacy. In *Nature and treatment of stuttering: new directions* (2nd ed., pp. 413-437).

Ingham, J.C. (2003). Evidence-based treatment of stuttering: I. Definition and application. *Journal of Fluency Disorders, 28,* 197-207.

Johnson, W. (1939). The treatment of stuttering. *Journal of Speech Disorders, 4,* 170-173.

Kent, R.D. (1984). Stuttering as a Temporal Programming Disorder. In R.F. Curlee & W.W.H. Perkins (Eds.), *Nature and Treatment of Stuttering: New Directions* (pp. 283-302). San Diego: College-Hill Press.

Kendall, D.L., McNeil, M.R., Shaiman, S., & Pratt, S. (2005). Phonetic encoding of infrequent articulatory phonetic transitions. Aphasiology, 19, 1, 39-52.

Kolk, H.J. (1991). Is stuttering a symptom of adaptation or of impairment? In H.F.M. Peters, W. Hulstijn, & C.W. Starkweather (Eds.), *Speech motor control and stuttering* (Amsterdam: Elseviers Science Publishers.

Kolk, H.J., & Postma, A. (1997). Stuttering as a covert-repair phenomenon. In R.F. Curlee & G.M. Siegel (Eds.), *Nature and treatment of stuttering: New directions* (2nd ed., pp. 182-203). Boston: Allyn & Bacon.

Lambert, M., & Barley, D. (2001). Lambert, M., & Barley, D. (2001). Research summary on the therapeutic relationship and psychotherapy outcome. *Psychotherapy, Theory, Research and Practice, 38,* 357-361.

Langevin, M., & Boberg, E. (1993). Results of an intensive stuttering therapy program. *Journal of Speech-Language Pathology and Audiology, 17,* 158-166.

Langevin, M., & Kully, D. (2003). Evidence-based treatment of stuttering: III. Evidence based practice in a clinical setting. *Journal of Fluency Disorders, 28,* 219-236.

Levelt, W.J.M. (1989). *Speaking: from Intention to Articulation.* Cambridge MA: MIT Press.

Levelt, W.J.M., & Wheeldon, L. (1994). Do speakers have access to a mental syllabary? *Cognition, 50,* 239-269.

Lukatela, G., Carello, C., Savic, M., UrosÏevic, Z., & Turvey, M.T. (1998). When nonwords activate semantics better than words. *Cognition, 68,* B31-B40.

McClean, M., Tasko, S.M., & Runyan, C.M. (2004). Orofacial movements associated with fluent speech in persons who stutter. *Journal of Speech, Language, and Hearing Research, 47,* 294-303.

Max, L. (2004). Stuttering and internal models for sensorimotor control: a theoretical perspective to generate testable hypotheses. In B. Maassen, R.D. Kent, H.F.M. Peters, P.H.H.M. Van Lieshout, & W. Hulstijn (Eds.), *Speech motor control in normal and disordered speech* (pp. 357-387). Oxford: University Press.

Miller, S., & Watson, B.C. (1992). The relationship between communication attitude, anxiety, and depression in stutters and nonstutterers. *Journal of Speech and Hearing Research, 35,* 789-798.

Onslow, M. (2003). Evidence-based treatment of stuttering: IV. Empowerment through evidence-based treatment practices. *Journal of Fluency Disorders, 28,* 237-245.

Packman, A., Onslow, M., Coombes, T., & Goodwin, A. (2001). Stuttering and lexical retrieval. *Clinical Linguistics and Phonetics, 15,* 487-498.

Peters, T., & Guitar, B. (1991). Stuttering An Integrated Approach to its Nature and Treatment. Baltimore, MD: Williams & Wilkins.

Postma, A., & Kolk, H.J. (1993). The covert repair hypothesis: Prearticulatory repair processes in normal and stuttered disfluencies. *Journal of Speech, Language, and Hearing Research, 36 (3),* 472-487.

Postma, A. (2000). Detection of errors during speech production: a review of speech monitoring models. *Cognition, 77,* 97-132.

Rastatter, M.P., & Dell, C.W. (1987a). Reaction time of moderate and severe stutterers to monaural verbal stimuli: some implications for neurolinguistic organization. *Journal of Speech and Hearing Research, 30,* 21-27.

Rastatter, M.P., & Dell, C. (1987b). Vocal reaction times of stuttering subjects to tachistoscopically presented concrete and abstract words: A closer look at cerebral dominance and language processing. *Journal of Speech and Hearing Research, 30,* 306-310.

Rastatter, M.P., & Dell, C.W. (1985). Reaction time of moderate and severe stutterers to monaural verbal stimuli: some implications for neurolinguistic organization. *Journal of Speech and Hearing Research, 30,* 21-27.

Schwartz, H.D., & Conture, E.G. (1988). Subgrouping young stutterers: preliminary behavioral observations. *Journal of Speech and Hearing Research, 31,* 62-71.

Seidenberg, M.S. & McClelland, J.L. (1989). A distributed, developmental model of word recognition and naming. *Psychological Review, 96,* 523-568.

Smith, A., & Goffman, L. (2004). Interaction of motor and language factors in the development of speech production. In B. Maassen, R. D. Kent, H. F. M. Peters, P. H. H. M. van Lieshout, & W. Hulstijn (Eds.), *Speech motor control in normal and disordered speech* (pp. 313-355). Oxford: University Press.

Stager, S.V., Denman, D.W., & Ludlow, C.L. (1997). Modifications in Aerodynamic Variables by Persons Who Stutter Under Fluency-Evoking Conditions. *Journal of Speech, Language, and Hearing Research, 40,* 832-847.

Stager, S.V., Calis, K., Grothe, D., Bloch, M., Berensen, N.M., Smith, P.J., & Braun, A. (2005). Treatment with medications affecting dopaminergic and serotonergic mechanisms: Effects on fluency and anxiety in persons who stutter. *Journal of Fluency Disorders, 30,* 319-335.

Starkweather, C.W., (1995). A simple theory of stuttering. *Journal of Fluency Disorders, 20,* 91-116.

Starkweather, C.W., & Gottwald, S.R. (1990). The Demands and Capacities model II: Clinical applications. *Journal of Fluency Disorders, 15,* 143-157.

Tasko, S.M., & McClean, M.D. (2004). Variations in articulatory movement with changes in speech task. *Journal of Speech, Language, and Hearing Research,* 47, 85-100.

Thomas, C., & Howell, P. (2001). Assessing efficacy of stuttering treatments. *Journal of Fluency Disorders, 26,* 311-333.

Tjaden, K. (1999). Can a model of overlapping gestures account for scanning speech patterns? *Journal of Speech, Language, and Hearing Research, 42,* 604-617.

Van Lieshout, P.H.H.M. (1995). *Motor planning and articulation in fluent speech of stutterers and nonstutterers.* dissertation University of Nijmegen.

Van Lieshout, P.H.H.M. (1997). Linguistic and motor determinants: how well can they be separated? In E.C. Healey & H.F.M. Peters (Eds.), San Francisco: Nijmegen University Press.

Van Lieshout, P.H.H.M., Peters, H.F.M., Starkweather, C.W., & Hulstijn, W. (1993). Physiological differences between stutterers and nonstutterers in perceptually fluent speech: EMG amplitude and duration. *Journal of Speech and Hearing Research, 36,* 55-63.

Van Lieshout, P.H.H.M., Hulstijn, W., & Peters, H.F.M. (2004). Searching for the weak link in the speech production chain of people who stutter: a motor skill approach. In B. Maassen, R.D. Kent, H.F.M. Peters, P.H.H.M. van Lieshout, & W. Hulstijn (Eds.), *Speech motor control in normal and disordered speech* (pp. 313-355). Oxford: University Press.

Van Lieshout, P.H.H.M., Hulstijn, W., & Peters, H.F.M. (1996). From Planning to Articulation in Speech Production: What Differentiates a Person Who Stutters From a Person Who Does Not Stutter? *Journal of Speech and Hearing Research, 39,* 546-564.

Van Lieshout, P.H.H.M., & Moussa, W. (2000). The assessment of speech motor behaviors using electromagnetic articulography. *The Phonetician, 81,* 9-21.

Van Riper, C. (1982). *The Nature of Stuttering.* (2nd ed.) Englewood Cliffs NJ: Prentice Hall.

Varley, R., & Whiteside, S.P. (2001). What is the underlying impairment in acquired apraxia of speech? *Aphasiology, 15,* 39-84.

Wilson, T.W., Leuthold, A.C., Lewis, S.M., Georgopoulos, A.P., & Pardo, P.J. (2005). The time and space of lexicality: a neuromagnetic view. *Experimental Brain Research, 162,* 1-13.

Wood, S.A.J. (1996). Assimilation or coarticulation? Evidence from the temporal co-ordination of tongue gestures for the palatalization of Bulgarian alveolar stops. *Journal of Phonetics, 24,* 139-164.

Wu, J.C., Maquire, G., Najafi, A., & Tang, C. (1997). PET Scan Evidence of Parallel Cerebral Systems Related to Treatment Effects: FDG and FDOPA pet scan Findings. In W. Hulstijn, H.F.M. Peters, & P.H.H.M. van Lieshout (Eds.), *Speech Production: Motor Control, Brain Research and Fluency Disorders* (pp. 329-341). Amsterdam: Elsevier.

Yaruss, J.S. (1998). Treatment Outcomes in Stuttering: Finding Value in Clinical Data. In A.K. Cordes & R.J. Ingham (Eds.), *Treatment Efficacy for Stuttering. A Search for Empirical Bases* (pp. 213-243). San Diego: Singular Publishing Group, Inc.

Yaruss, J.S. (1998). Describing the consequences of Disorders: Stuttering and the International Classification of Impairments, Disabilities, and Handicaps. *Journal of Speech, Language, and Hearing Research, 41,* 249-257.

Yaruss, J.S. (2001). Evaluating treatment outcomes for adults who stutter. *Journal of Communication Disorders, 34,* 163-182.

Yaruss, J.S., & Quesal, R.W. (2004). Stuttering and the international classification of functioning, disability and health (ICF): an update. *Journal of Communication Disorders, 37,* 35-52.

Ziegler, W. (2001). Apraxia of speech is not a lexical disorder. *Aphasiology, 15,* 74-77.

Ziegler, W. (2003). Speech motor control is task-specific: evidence from dysarthria and apraxia of speech. *Aphasiology, 17,* 3-36.

Zimmermann, G. (1980). Articulatory Dynamics of Fluent Utterances of Stutterers and nonstutterers. *Journal of Speech and Hearing Research, 23,* 95-107.

Summary

In this thesis, two aspects of stuttering are studied. The first part focuses on treatment efficacy and the second part on aspects of speech motor control in stuttering and non-stuttering persons. Both parts are related because therapy programs are based on ideas about the cause of stuttering and, in turn, treatment efficacy information importantly contributes to our understanding of stuttering as a complex phenomenon.

Part I. Stuttering therapy efficacy

The first part of this thesis (Chapter 2, 3, 4, and 5) focuses on treatment efficacy. Chapter 2 describes the results of a longitudinal efficacy study that was initiated in January 1999 and completed at the end of 2003. In this study the effects of the following three stuttering therapy programs were evaluated: (1) The *'Comprehensive Stuttering Programme'* (CSP), an integrated treatment program based on fluency-enhancing techniques, tension and stuttering modification, and on cognitive behaviour strategies to help the client deal with the emotional and attitudinal aspects of stuttering. Within the CSP, the relative weights of the components may vary across clients, and, compared to the other two programmes that were investigated, more emphasis (in terms of time) is put on speech-motor control. (2) The '*Doetinchem Method*' (DM), a group therapy that focuses on the social perspective of stuttering, *i.e.* the stutterer's social context is the starting-point for the therapy. This programme particularly targets the emotional and cognitive components of stuttering but also tries to enhance speech fluency. (3) The individualised therapy according to the protocol of the '*Vereniging Stottercentra Nederland*' (VSN), this is the Dutch Accociation of Stuttering Therapy Centres. This individual stuttering intervention is aimed at the client's specific speech-motor components and the stuttering-related emotional and cognitive aspects and is founded on a diagnostic procedure prior to the start of therapy. All participants were tested immediately before the start of therapy (pretreatment assessment), immediately after the end of the therapy (post-treatment assessment), one year post therapy (follow-up 1), and two years post therapy (follow-up 2). In each test session, the participants' speech was studied at two levels of observation: (1) s*peech behaviour,* as evaluated by independent raters and (2) assessments based on *self-reports*. Speech behaviour was studied by measuring stutter severity, speech quality and speech-motor control. Stutter severity was assessed in terms of percentage of stuttered syllables (%SS), syllables per minute (SPM) and self-evaluation (or speech satisfaction) on a rating scale from one to ten. Speech quality was measured using a perceptual rating instrument to judge speech characteristics (speech rate, naturalness, intonation, etc.) at each of the four assessment sessions. Speech-motor control was studied using the diadochokinesis (DKK) task derived from the Nijmegen Speech Motor Test (NSMT). With respect to the self-reports, the participants completed several questionnaires about stutter severity, speech anxiety and attitudes. The results are summarised next.

The most striking result of the CSP programme was the spectacular decrease in the number of disfluencies. Our participants spoke with almost complete fluency at the post-treatment evaluation. This effect was in line with the aim and content of the programme in that more than 70% of the time was spent on speech-motor control. Speech rate increased as a result of the drop in disfluencies. The superior post-treatment DDK results indicated an overall improved speech capacity. This effect was

maintained in the long term. Furthermore, although the clients showed some regression after the post-treatment assessment, relative to the pretreatment levels the treatment effects were sustained. The most pronounced relapse (regression) was found in the number of disfluencies, and this result was reflected in the speech satisfaction ratings. Nevertheless, the participants' own long-term speech ratings were significantly above the pretreatment level indicating that, despite the slight relapse, they were satisfied with the results achieved. Moreover, the long-term articulation quality improved, as was shown in the two follow-up assessments. The scores on the self-report questionnaires showed a decrease in stutter severity and speech anxiety and improved attitudes for both the intermediate and long-term follow-up.

The results of the DM were characterised by the enhanced self-perception scores immediately after the end of therapy. Moreover, the gains were rather robust and were maintained in the long term. This is in line with the DM focus on the reduction of negative factors that maintain stuttering. The gains were highest for the speech satisfaction scale and for the self-report questionnaires. There was hardly any relapse in the DM group, in contrast to the significant relapse observed in the CSP group. Compared to the CSP there were fewer improvements on speech fluency. However, it should be noted that at the pretreatment measurement the DM group showed fewer disfluencies compared to the CSP group and thus, there was less to be gained. Unfortunately, the moderate post-treatment decrease in disfluency was not maintained in the long term. On the other hand, the ratings for voice dynamics and speech power factors had improved. This means that the speech in the DM group sounded more dynamical and more powerful and, again, these effects appeared to be rather robust. There was no intermediate or long-term relapse for either factor. The absence of any effect on the articulation quality factor was in line with the fact that the DM mainly focuses on emotions and cognitions that are related to the stuttering problem and less treatment time is allocated to the enhancement of speech fluency and articulation quality.

The most prominent result in the VSN group was the improvement in the self-perception scores, in particular on the speech satisfaction scale. Although the mean post-treatment percentage of stuttered syllables was reduced, the difference was not significant. This may be due to the small sample size, the relatively large standard deviation but also to the smaller reduction in post-treatment disfluency (compared to CSP). Speech satisfaction had improved at post-treatment evaluation but in the long term there was a small relapse. Thus, even though no real relapse at the level of disfluency was found, speech satisfaction decreased in the long term. At the post therapy evaluations the judgments on the voice dynamics factor, the articulation factor and the speech power factor had not improved. The effects did not reveal themselves until the follow-up sessions. The voice dynamics were rated higher in the second follow-up and augmented speech power scores were reported in the two follow-up assessments. A similar pattern was found for the DDK task: there were no post-treatment or intermediate follow-up effects but in the 2-year follow-up the articulation capacity on the /pətəkə/ sequence was judged as having improved.

In conclusion, each of the three therapy programmes evaluated were found to be effective, but regression with respect to the percentage of stuttered syllables and also with respect to speech satisfaction proved to be inevitable. Based on these results we recommended independent diagnostic examinations prior to therapy that assess the client-specific physical, cognitive and emotional aspects of stuttering. Individual and group therapies offer different possibilities, which implies in our view that the inclusion of various approaches within one organisation, as is the case with the VSN and DM programmes in the integrated care programme (the so-called Integraal Zorg Traject or IZT), will offer superior treatment outcomes for each individual client.

After having evaluating the effect of these therapy programs, we examined if different profiles of stuttering were sensitive to different treatment approaches. In the next study (Chapter 3), we first explored a procedure for sub-typing individuals who stutter, followed by a study of the relationship between sub-typing and treatment outcome. Twenty-five adult participants of the CSP were classified according to (1) stuttering severity and (2) severity of negative emotions and cognitions associated with their speech problem. Speech characteristics (percentage of stuttered syllables, distorted speech score, and the number of correctly produced syllables on a DDK task) and emotional/cognitive states (emotional reaction, speech satisfaction, and attitudes toward speaking) were assessed before and after treatment, and at a 1- and 2-year follow-up. The results showed that (a) there was no relationship between stuttering severity and the severity of negative emotions and cognitions, (b) the severe stuttering group had the largest treatment gains but also the highest level of regression, and (c) at post-treatment and both follow-up assessments the differences on measures of emotions between the mild and severe emotional group had disappeared, chiefly due to a large decrease in the latter group's negative emotions and cognitions. Our findings showed that, based on treatment gains, specific subgroups can be identified. This suggested that different treatment approaches are required and underlines the necessity of developing a better understanding of how various dimensions of stuttering relate to treatment outcome.

The CSP programme, developed at the Institute for Stuttering Treatment and Research (ISTAR) in Canada, showed to be quiet effective in Canada. In Chapter 4 it was examined if these results could be generalized from one country (Canada) to another (The Netherlands). We showed that both Canadian and Dutch participants were maintaining clinically meaningful reductions in stuttering frequency (in within and beyond clinic measures) and improvements in self-report measures of communication attitudes, confidence, and perceptions of stuttering. Treatment effect sizes across all measures ranged from medium or typical to larger than typical. With the exception of speech rate measures, there were no significant differences between the Dutch group and the Canadian comparison group in stuttering frequency in the beyond clinic condition or self-report measures. Effect sizes for the pooled Dutch and Canadian data ranged from typical to larger than typical across stuttering frequency and self-report measures. The most notable difference between the groups occurred in participants' self-perceptions. Almost half of Dutch participants indicated that they no longer considered themselves a stutterer at 2 years follow-up. These data suggest that there may be cultural differences in self-perceptions after treatment. However, replication of these results is needed before implications for treatment, treatment outcome, and future research can be determined.

In the initial efficacy study (Chapter 2), a large (and time-consuming) battery of assessment tools was used in the evaluation of treatment efficacy. In Chapter 5, the validity of a simple and not time-consuming self-assessment scale (School Marks: SM) was tested by relating it to (1) objective measures (percentage stuttered syllables: %SS and syllables per minute: SPM) and (2) the results of (self-)evaluation tests (questionnaires and perceptual evaluations or judgements of disfluency and naturalness by naïve listeners). Data were from the CSP and the DM group at pretherapy, post therapy and one and two year follow up were selected. Results showed similar score patterns on the SM assessment, the self-evaluation questionnaires and the objective measures over time. We concluded that the validity of the SM measure was proved. Therefore we encourage the use of such an instrument when (stuttering) treatment efficacy is studied.

Part II. Linguistic and motor factors in stuttering

The second part of this thesis focuses on aspects of speech motor control in person who stutter (PWS) and persons who do not stutter (PWNS). In the articles of part II, subscales of 'The Nijmegen Speech Motor Test' (NSMT) were analysed. This test was developed to classify participants into specific stuttering profiles, based on aspects of speech motor control. Better understanding of stuttering profiles results in better therapy choices, resulting in improved treatment outcomes. In the first part of this thesis, classification of stuttering was based on stuttering symptoms on the surface of stuttering (*e.g.* speech rate, percentage stuttered syllables, struggle behaviour). With the NSMT, *underlying* aspects of stuttering were measured, using a simultaneous registration of phonation, articulation and respiration in a set of speech tasks. In other words, the NSMT was designed to assess underlying timings aspects in the speech motor processes. Part II of this thesis describes the first results of several speech tasks of the NSMT.

In Chapter 6, we studied if PWS differ from PWNS in the coproduction (also coarticulation or blending) of different types of consonant clusters. In former studies (see Tjaden, 1999 for a review) it was shown that articulatory deficits in motor speech disorders often are explained in the light of overlapping articulatory gestures. Based on the Gestural Phonology Model of Browman and Goldstein, two types of consonant clusters were formed: homorganic and heterorganic clusters, both intra-syllabic (CVCC) and inter-syllabic (CVC#CVC). In homorganic clusters the consonants have the same place of articulation (*i.e.*, belong to the same tier) whereas in heterorganic clusters the consonants have different places of articulation (*i.e.*, belong to different tiers). Differences between the two groups were measured in the number of disfluencies, incorrect produced speech productions, in speech reaction times and in word durations. It was predicted to: (1) Find longer RT's in the homorganic condition compared to the heterorganic condition. This was based on the notion that more time is needed for the planning and/or initiation of sequential movements from the same articulatory tier (2) Find shorter WD's in the heterorganic condition. This was based on the Gestural Phonology model, which predicts more overlap and thus less execution time in heterorganic clusters. (3) Find similar effects in both groups, but only to a stronger extent in the PWS group. In former studies it was shown that PWS differ from PWNS in the time to complete speech production. (4) Find a stronger WD effect of cluster type in the CVCC condition. This was based on the assumption that restrictions on coarticulation are less within a syllable than across a syllable boundary.

Results showed that homorganic clusters elicited more incorrect speech productions and longer reaction times than the heterorganic clusters, but there was no difference between the homorganic and the heterorganic clusters in the word duration data. Persons who stutter showed a higher percentage dysfluencies and a higher percentage incorrect speech productions than PWNS but there were no main group effects in reaction times and word durations. However, there was a significant three-way interaction effect between group, cluster type and cluster place: homorganic clusters elicited longer reaction times than heterorganic clusters, but only in the inter-syllabic condition and only for persons who stutter. These results suggest that the production of two consonants with the same place of articulation across a syllable boundary puts higher demands on motor planning and/or initiation than producing the same cluster at the end of a syllable, in particular for PWS.

Chapter 7 was set up to find relevant tasks to discriminate between PWS and PWNS and to show effects of manipulations on both groups. Differences between these two groups were studied on word length, word meaning, sentence repetition, consonant cluster place and DDK. Acoustic reaction times and word durations were measured. Overall, PWS showed significantly longer reaction times and word durations than PWNS. Only in the reaction time data of the word meaning task the interaction between group and the main factor word versus non-word was significant. Furthermore, the largest percentage difference between the two groups was found in the Mean Absolute Deviation of the utterance duration in the sentence repetition task: PWS showed more variation in their speech production than PWNS.

In Chapter 8 the results of these studies are discussed. We attempted to bridge the gap between theories of stuttering and the practice of therapeutic interventions. Our theory showed that people that stutter lack the mental recourses to remain fluent in complex speech tasks, and, as a result, need more time and mental capacities to remain fluent. Speech techniques acquired in stuttering therapy seem to improve these timing aspects of stuttering. However, efficacy is related to different types of stuttering profiles and therapy programmes and diagnostic procedures should be adapted to fit these profiles.

Samenvatting

Dit proefschrift beschrijft twee aspecten van stotteren: (1) het effect van stottertherapieën en (2) spraakmotorische controle bij stotteraars en bij niet-stotteraars. Beide aspecten zijn aan elkaar gerelateerd omdat therapieprogramma's gebaseerd zijn op aannames over de oorzaak van stotteren en, aan de andere kant, informatie met betrekking tot de werking van therapie op belangrijke wijze bijdraagt aan onze kennis over het complexe fenomeen stotteren.

Deel I. De evaluatie van stottertherapie

In hoofdstuk 2 worden de resultaten van een uitgebreid longitudinaal effectonderzoek (van januari 1999 tot eind 2003) beschreven. In dit onderzoek worden de effecten geëvalueerd van: (1) Het "*Comprehensive Stuttering Program*" (CSP), een geïntegreerd behandelplan waarin naast het aanleren van vloeiende spraak, aandacht wordt besteed aan de emotionele en cognitieve componenten van het stotteren. Deze therapie wordt in groepsverband gegeven. Binnen het therapieprotocol varieert de aandacht die aan deze componenten besteed wordt per individu. (2) De *"Doetinchemse Methode"* (DM), een groepstherapie waarbij het stotteren vanuit een sociale context wordt benaderd. Het verminderen van negatieve factoren die het stotteren in stand houden en het opbouwen van vloeiende spraak staan centraal. (3) Individuele therapie die wordt gegeven volgens het protocol van de *"Vereniging Stottercentra Nederland"* (VSN). Deze therapie stelt op basis van een diagnostisch onderzoek een individueel therapieprogramma samen, waarbij de verdeling van de spraakmotorische, emotionele en cognitieve componenten van de therapie wordt afgestemd op de problematiek van de individuele stotteraar.

De deelnemers werden direct vóór therapie, direct na therapie, één jaar na therapie (Follow up 1; F1) en twee jaar na therapie (Follow up 2; F2) getest. Het spraakgedrag werd bestudeerd door het meten van stotterernst, analyse van de spraakkwaliteit en het meten van de spraakmotoriek. Voor de bepaling van de stotterernst werden het percentage gestotterde syllaben (%SS), het aantal syllaben per minuut (SPM) en zelfevaluatie met behulp van rapportcijfers bepaald. De spraakkwaliteit werd onderzocht met behulp van een beoordelingsinstrument waarin spraakkenmerken zoals spreeksnelheid, natuurlijkheid, intonatie etc. beoordeeld werden. De spraakmotoriek werd onderzocht met behulp van de diadochokinesetaak (DDK-taak) uit de Nijmeegse Spraakmotoriek Test (NSMT). Om inzicht te krijgen in de zelfbeleving van de stotteraar werd gebruik gemaakt van een aantal bestaande vragenlijsten waarin de zelfbeleving met betrekking tot stotterernst, spreekangst en attitude geïnventariseerd werd. De resultaten kunnen als volgt worden samengevat.

Het CSP. Direct na therapie werd de spraak vloeiender, de spreeksnelheid nam toe en werd er hoger gescoord op de zelfevaluatieschaal. Ook gaven naïeve (d.w.z. hiervoor niet getrainde) luisteraars een verbetering in de articulatiekwaliteit aan. Er werd een sterke vooruitgang weergegeven in de vragenlijsten die betrekking hadden op stotterernst, spreekangst, attitude en sociale angst. In de spraakmotoriek werd een hoger aantal correct geproduceerde syllaben in de /pətəkə/-reeks van de DDK-taak gevonden. Ondanks een terugval ten opzichte van de situatie direct na therapie, werden op F1 en F2 belangrijke verbeteringen behouden in de vloeiendheid, de spreeksnelheid, de zelfevaluatie, de articulatiekwaliteit, de beleving van stotterernst, de spreekangst, de attitude en de sociale angst. Het aantal correct geproduceerde syllabes in /pətəkə/ bleef ook in F1 toenemen. Drieëntachtig procent van de deelnemers gaf aan dat de CSP hen zeer goed of redelijk goed had geholpen.

De DM. Direct na therapie werd de spraak vloeiender en steeg de spreeksnelheid in de monoloog. De zelfevaluatie van de stotteraars werd positiever. De therapie had een positief effect op de beleving van het stotteren. Naïeve luisteraars namen een verbeterde stemdynamiek en een toegenomen draagkracht van de spraak waar. Met betrekking tot de spraakmotoriek werd een hoger aantal correct geproduceerde syllaben in de /pətəkə/-reeks van de DDK-taak gevonden. Op F1 en F2 was er een kleine terugval ten opzichte van de situatie direct na therapie. Een positief therapie-effect bleef behouden op het gebied van spreeksnelheid en zelfevaluatie, maar er was geen effect meer op de vloeiendheid van de spraak. De stemdynamiek en draagkracht van de spraak, maar ook grotendeels de beleving van het stotteren bleven ook op lange termijn verbeterd. De maximale spraakproductiescore in de syllabereeks /pətəkə/ bleef toenemen. Meer dan 80% van de deelnemers gaf aan dat de DM hen zeer goed of redelijk goed had geholpen.

De VSN. Direct na therapie was de spreeksnelheid toegenomen en was de zelf-evaluatie positiever geworden. Er werd geen effect op de stemdynamiek, de articulatie-kwaliteit en de draagkracht van de spraak aangetoond en er werd geen significante toename van het aantal correct geproduceerde syllaben in de /pətəkə/-reeks van de DDK-taak gevonden. Er was een kleine vooruitgang in de beleving van stotterernst, spreekangst en aspecten van sociale angst. Op F1 en F2 vonden we een kleine terugval, maar de significante verbeteringen met betrekking tot de zelfevaluatie bleven behouden en de draagkracht van de stem was verbeterd. Het aantal correct uitgesproken syllaben in de syllabereeks /pətəkə/ van de DDK-taak was significant hoger dan vóór therapie. De verbetering in de beleving van stotterernst bleef behouden. Meer dan 80% van de deelnemers gaf aan dat de VSN zeer goed of redelijk goed had geholpen.

Uit bovenstaande resultaten werd geconcludeerd dat de drie betrokken therapieën op korte en lange termijn effectief zijn. Omdat elke behandelvorm zijn eigen kwaliteiten heeft is het van belang dat er, voorafgaand aan de therapiekeuze onafhankelijk diagnostisch onderzoek plaatsvindt, zodat op basis hiervan de meest geschikte therapievorm gekozen kan worden. Hierbij werd ervan uitgegaan dat stotteraars met verschillende stotterprofielen anders reageren op verschillende behandelmethoden. Om dit aan te tonen wordt in Hoofdstuk 3 een methode voor het indelen van stotteraars in verschillende subtypen voorgesteld. Op basis van deze indeling werd de relatie tussen deze subtypen en de therapie-effecten onderzocht. Vijfentwintig CSP-deelnemers werden in dit onderzoek ingedeeld op basis van: (1) stotterernst (ernstig stotteren: ES; matig stotteren: MS) en (2) de ernst van met spreken/stotteren samenhangende negatieve emoties en cognities (ernstig emotioneel: EE; matig emotioneel: ME). Spraakkenmerken (percentage gestotterde syllaben, mate van afwijkende spraak en attitude ten opzicht van spreken) werden onderzocht vóór therapie, na therapie, op F1 en F2. De resultaten toonden aan dat (a) stotterernst en de ernst van negatieve emoties en cognities niet met elkaar samenhangen, (b) de groep geclassificeerd als 'ernstige stotteraars' de grootste therapiewinst behaalt maar ook de grootste terugval vertoont, en (c) direct na therapie en bij F1 en F2 de gevonden verschillen tussen de EE en ME groepen verdwenen zijn, voornamelijk als gevolg van een sterke afname van negatieve emoties en cognities in de EE groep. De data toonden ook aan dat op basis van therapie-effecten specifieke subgroepen onderscheiden kunnen worden. Hieruit werd geconcludeerd dat verschillende behandelprogramma's noodzakelijk zijn en dat het van belang is beter inzicht te krijgen in hoe de verschillende dimensies van stotteren gerelateerd zijn aan therapie-effecten.

Van de CSP was al eerder aangetoond dat deze in Canada effectief is. In Hoofdstuk 4 is onderzocht of deze Canadese resultaten gegeneraliseerd kunnen worden naar de Nederlandse populatie. In deze studie werd zowel in de Canadese als de Nederlandse groep een belangrijke afname in onvloeiendheden gevonden. Bovendien werd in beide groepen vooruitgang geboekt in zelfevaluatie, communicatieattitude, zelfverzekerdheid en de waarneming van het stotteren. Met uitzondering van de resultaten bij spreeksnelheid waren er geen significante verschillen tussen de Nederlandse en de Canadese groep.

De in hoofdstuk 2 beschreven therapie-evaluatie werd uitgevoerd met behulp van een grote (en hierdoor tijdrovende) testbatterij. In hoofdstuk 5 wordt de validiteit van een eenvoudige en weinig tijdrovende zelfevaluatieschaal onderzocht. In dit onderzoek werd deze zelfevaluatieschaal gerelateerd aan (1) objectieve maten (percentage gestotterde syllaben en het aantal uitgesproken syllaben per minuut, en (2) de resultaten van een reeks bestaande (zelf)evaluatietesten (o.a. vragenlijsten en beoordelingsschalen). De data van de twee geïncludeerde groepen (De DM groep en de CSP groep) op vier tijdstippen (vóór therapie, na therapie, F1 en F2) toonden overeenkomstige scorepatronen (in de tijd) met de onderzochte zelfevaluatiemaat, de bestaande vragenlijsten en de objectieve maten. Deze resultaten toonden aan dat de eenvoudige zelfevaluatieschaal een valide maat is voor het meten van therapie-effecten.

Deel II. Linguïstische en spraakmotorische factoren in stotteren

In het tweede deel van dit proefschrift worden verschillende aspecten van de spraakmotorsiche controle onderzocht in een vergelijking van groepen stotteraars (ST) en niet-stotteraars (NST). Hiertoe werd gebruik gemaakt van subschalen van de NSMT. Deze test werd ontwikkeld om de deelnemers in stotterspecifieke profielen in te delen op basis van spraakmotorische controle. Meer inzicht in stotterprofielen leidt tot betere therapiebeslissingen, wat zal resulteren in betere uitkomsten. In tegenstelling tot het eerste deel van dit proefschrift, waarin de deelnemers ingedeeld werden op basis van uitwendig waarneembaar stottergedrag (zoals spreeksnelheid, percentage gestotterde syllaben en vechtgedrag), werden in deel II onderliggende aspecten van de spraak gemeten. Er werd gebruik gemaakt van de NSMT waarmee, door simultane registratie van fonatie, articulatie en respiratie in een set spreektaken, timings- en plannings-aspecten in spraakmotorische processen onderzocht kunnen worden. Van een aantal van deze taken worden de eerste resultaten in deel II van het proefschrift gepresenteerd.

Hoofdstuk 6 handelt over de verschillen tussen ST en NST op het gebied van coarticulatie van verschillende soorten consonantclusters. In eerdere studies is aangetoond dat de articulatorische problemen in aandoeningen van het spraakmotorische systeem vaak verklaard kunnen worden aan de hand van 'overlappende articulatie-bewegingen' (gestural overlap). Op basis van het zogenaamde 'Gestural Phonology Model' (GPM) van Browman & Goldstein zijn twee typen consonantclusters te onderscheiden: (a) consonantclusters waarbij beide consonanten dezelfde plaats van articulatie hebben (zogenaamde homorganic (HO) clusters) en (b) clusters waarbij de consonanten verschillende articulatieplaatsen hebben (zogenaamde heterorganic (HE) clusters). Beide typen cluster werden zowel intra-syllabisch (CVCC) als inter-syllabisch (CVC#CVC) aangeboden (C=consonant, V=vocaal). Verschillen tussen beide groepen werden gemeten in het aantal onvloeiendheden, het aantal foutieve spraakproducties, en in spraakreactietijden en woorddduren. Op basis van het GPM werd het volgende voorspeld: (1) Er is meer tijd nodig voor het plannen en/of initiëren van HO clusters. Dit was gebaseerd op de aanname dat er meer tijd nodig is voor het plannen en/of initiëren van opeenvolgende bewegingen van dezelfde articulator. (2) De woordduren zijn korter in de HE conditie. Dit was gebaseerd op de voorspelling van het GPM dat er meer overlap mogelijk is en daardoor minder executietijd nodig is in de heteroganic clusters. (3) Beide groepen vertonen dezelfde effecten, maar deze effecten zijn in de ST groep in sterkere mate aanwezig. (4) De reactietijdeffecten zijn geprononceerder in de CVCC conditie. Dit was gebaseerd op de aanname dat coarticulatie binnen een syllabe sterker is dan over een syllabegrens. De resultaten toonden aan dat HO clusters meer onjuist geproduceerde spraakproducties uitlokten in vergelijking met HE clusters. Er werd geen verschil tussen de twee typen clusters in woordduren gevonden. De ST vertoonden (vanzelfsprekend) meer onvloeiendheden dan de NST maar daarnaast ook een hoger percentage onjuist geproduceerde spraakproducties. Er was geen groeps-effect in de reactietijden en in woordduren. Wel werd er een significante drie-weg-interactie gevonden tussen groep, cluster type en cluster plaats: HO clusters hadden langere reactietijden dan HE clusters, zij het alleen in de inter-syllabische conditie en alleen in de ST-groep. Deze resultaten doen vermoeden dat vooral in de ST groep de productie van twee consonanten met dezelfde articulatieplaats in de CVC#CVC conditie meer 'eisen' van de motorische planning en/of initiatie dan de productie van dit cluster in de coda positie van een syllabe.

In Hoofdstuk 7 wordt een onderzoek beschreven waarin we relevante taken die kunnen discrimineren tussen ST en NST, evenals de effecten van verschillende manipulaties (woordlengte, woordbetekenis, zinherhaling, plaats van het consonant-cluster en in DDK-vaardigheid) op beide groepen. De afhankelijke variabelen waren akoestische reactietijden en woordduren. De ST groep had significant langere reactietijden en woordduren dan de NST groep. In de woordbetekenistaak was het verschil in reactietijd tussen de ST groep en de NST groep groter bij nonwoorden dan bij woorden; deze interactie was significant. Daarnaast hadden ST meer variatie in hun spraakproductie bij de zinherhalingtaak; het grootste verschil tussen de twee groepen werd gevonden in de Mean Absolute Deviation van de zinsduren in deze taak.

In Hoofdstuk 8 worden de resultaten van bovengenoemde studies nader besproken. Hierin pogen wij een brug te slaan tussen de theorieën over stotteren en de gehanteerde therapeutische interventies. Uit onze theorie bleek dat personen die stotteren bepaalde mentale vermogens missen om vloeiend in complexe spraaktaken te blijven. Als gevolg hiervan hebben ze meer tijd en mentale capaciteit nodig om vloeiend te blijven spreken. De spraaktechnieken zoals die worden geleerd in stottertherapie lijken dit soort timingaspecten van het stotteren te verbeteren. Hierbij dient echter aangetekend te worden dat de therapieresultaten gerelateerd zijn aan het stotterprofiel en het therapieprogramma. Daarom zouden therapieprogramma's aangepast moeten worden aan deze profielen.

Dankwoord

Graag wil ik alle mensen bedanken die me de afgelopen jaren direct of indirect hebben geholpen bij het tot stand brengen van dit proefschrift. In het bijzonder dank ik:

Mijn promotoren *Prof. dr. W. Hulstijn* en *Prof. dr. K. Graamans*. Beste Wouter, de uitvoerige analyses en discussies met betrekking tot de reactietijddata waren leerzaam en boeiend. We hebben heel wat gebrainstormd over de mogelijke interpretaties. Hartelijk dank hiervoor. Beste Professor Graamans, ondanks de afstand van dit onderwerp tot uw vakgebied, las u de teksten met interesse en waren deze altijd binnen twee weken voorzien van zinvol commentaar of, zoals u het noemde, suggesties. Hartelijk bedankt hiervoor en vooral ook voor uw steun in de laatste fase van dit traject.

Mijn co-promotoren*, dr. H. Peters* en *dr. B. Maassen*. Beste Herman, als initiator en inspirator van het evaluatieonderzoek ben je vanaf het begin tot het einde bij het onderzoekstraject betrokken geweest. Je wees me de weg in de wereld van het stotteren en stimuleerde me om me hierin te verdiepen, congressen te bezoeken en teksten te schrijven. Zelfs na je afscheid kwam je nog één keer per week terug, onder andere om dit traject mee af te kunnen ronden. Dat heb ik zeer gewaardeerd. Bedankt voor je steun en vertrouwen in mij! Beste Ben, pas in een later stadium van dit traject deed ik een beroep op je expertise. Dit leidde tot een aantal boeiende inhoudelijke discussies. Je efficiënte benadering in de laatste fase van het traject gaf steun en houvast. Bedankt hiervoor!

Zeer veel dank gaat naar *dr. P. Van Lieshout*. Beste Pascal, vanuit Canada dacht je mee. Jouw kennis over STP-gerelateerde theorieën is indrukwekkend. Elke vraag die ik je stelde werd snel en uitvoerig beantwoord. Ik dank je dan ook voor deze prettige samenwerking en alle interessante en leerzame discussies!

Ook heel veel dank gaat uit naar *Prof. dr. T. Rietveld*. Beste Toni, het was een genoegen om met je samen te werken. Ik hoop dat nog vele gezamenlijke artikelen mogen volgen.

I would like to thank *Marilyn Langevin* and *Deborah Kully* for the introduction of the CSP in the Netherlands and for the fruitful collaboration on the articles. I commend you both for your care, integrity and professionalism. Marilyn, you visited the Netherlands several times but most of our contact was through e-mail. Your move from Canada to Australia was also a move from practitioner to scientist, which made our contact more frequent. It was great working together with you. I think we understood each other very well and that we share a similar drive and spirit in our work! Thank you Marilyn, finally we both did it!

Esther Wouters, jou wil ik bedanken voor de zeer prettige samenwerking. Als onderzoeksassistent heb je heel wat werk verricht en met je nuchtere houding ten aanzien van wetenschappelijk werk 'floot' je mij (onbewust denk ik) zo nu en dan eens terug. Je was een leuke kamergenoot.

Harrydath Gisaidobee, Esther Wittenburg, Monica Tognon, 'counters': *Anke, Anette, Judith & Maaike*, hartelijk bedankt voor de enorme inzet in de eerste fase van dit traject.

Studenten van de opleiding Spraak-Taalpathologie *Judith Verberne, Heidi, Diepstra, Joyce van Daalen, Pleuntje Molkenboer en Ans Kilkens* en de studenten van de opleiding logopedie dank ik voor het werk en de inzet. Het was leuk te zien dat jullie mijn enthousiasme voor het vak deelden. Jullie vragen maakten me wijzer.

Ruud Bökkerink, jij was vanuit de opleiding logopedie van de Hogeschool van Arnhem en Nijmegen betrokken bij de introductie van de CSP in Nederland. Dit was erg succesvol. Het was prettig met je samen te werken. *Marian Tellers*, jouw expertise op het gebied van stotteren in het algemeen en de CSP in het bijzonder in combinatie met het vele werk dat je hebt verricht, maakte de introductie van de CSP in Nederland succesvol. Bedankt voor de prettige samenwerking.

Alle therapeuten van de VSN, CSP en DM en de mensen van de klankbordgroep dank ik voor hun medewerking. De open houding ten aanzien van evaluatieonderzoek sterkt jullie methode.

Natuurlijk gaat veel dank uit naar alle mensen die participeerden in het onderzoek. Heel wat uren kwamen jullie naar onze afdeling voor allerlei metingen en onderzoeken. Zonder jullie participatie was dit onderzoek nooit tot stand gekomen. Ook de studenten van de Hogeschool van Arnhem en Nijmegen die meededen aan het luister-experiment dank ik voor hun inzet.

Mijn lieve collega's van de 'afdeling' Stem- en spraak: *Bertie, Monique, Felix, Piet, Mariëtte, Rosemary (jij in het bijzonder), Tiny en Anneke(†).* Toen ik door reorganisatie met het lab moest verhuizen naar de researchgang realiseerde ik me hoe veel jullie voor me hebben betekend als directe collega's! Bedankt hiervoor!

Gelukkig werd ik goed opgevangen door mijn nieuwe gang/borrelgenoten *Joop, Aukje, Carlo, Bart, Chris-Jan, Jan, Christoph, Lucas*. Mijn eenzame wetenschappelijke worstelingen werden door jullie op zeer welkome wijze onderbroken in de pauze. Samen lunchen, wandelend in de zon, op een bankje of op het gras. Bedankt hiervoor!

Lian Nijland, in de laatste fase van het traject was jij mijn kamergenoot. We hebben veel wetenschap en niet-wetenschap gedeeld! Bedankt hiervoor.

Ook wil ik alle leden van de Werkgroep SSTP hartelijk bedanken voor de wetenschappelijke input tijdens het (min of meer) maandelijkse overleg.

Andy Beynon, jij stimuleerde me om nu eindelijk dit traject eens af te ronden. Bedankt, je had gelijk!

Mijn nieuwe collega's van het Cochleaire Implantatie (CI) team, en in het bijzonder "collega-coördinator" *Esther Dekkers* (ik kon het niet beter treffen), bedankt voor jullie steun bij het afronden van het proefschrift. Het secretariaat van het CI-team

Anne-Marie, Sabine, Iris en Arzu bedankt voor jullie begrip voor al die stapels papieren op mijn bureau… binnenkort ga ik opruimen!

Edith Huinck en *Mariette Embrechts,* bedankt dat jullie mijn paranimfen willen zijn.

Marnix Donders en *Loe van Olffen* dank ik voor alle handige tips om binnen een relatief korte tijd een omslag voor dit proefschrift te kunnen maken. Bedankt *Hanneke Meulenbroek* voor het corrigeren van het Engels en *Diny Helsper* voor de hulp bij de lay-out.

Natuurlijk dank ik al mijn lieve vrienden, vriendinnen en (schoon)familie voor de steun die ik kreeg. We zagen elkaar erg weinig de laatste tijd… ik ga mijn leven beteren! Mijn zangvriend(inn)en: Bedankt voor de altijd zeer welkome muzikale afleiding op donderdagavond en… op naar ons eerste optreden!

Mijn lieve ouders, het is geweldig jullie op deze wijze te kunnen bedanken voor jullie onvoorwaardelijke steun en voor alles wat jullie me hebben meegegeven om dit te kunnen bereiken!

Lieve Rogier, je leerde me veel door al onze discussies over wetenschap in het algemeen en statistiek in het bijzonder. Bedankt voor je begrip voor mijn gedrevenheid en voor alle steun die je me gaf. Jij, Bas en Sem zijn het beste wat me ooit is overkomen!

Publications

Hartsuiker, R.J., Kolk, H.H.J., & **Huinck, W.J.** (1998). Subject-verb agreement construction in agrammatic aphasia. *Brain and Language, 65* (1), 64-67.

Hartsuiker, R.J., Kolk, H.H.J., & **Huinck, W.J.** (1999). Agrammatic Production of Subject-Verb Agreement: The Effect of Conceptual Number. *Brain & Language 69,* 119-160.

Huinck, W.J., & Hartsuiker, R.J. (1997). One or more labels on the bottles? Production and comprehension of distributed phrases in Broca's aphasia. *Abstract book Nederlandse Vereniging voor Psychonomie 1997.*

Huinck, W.J., Langereis, M.C., Lamoré, P.J.J., & Smoorenburg, G.F. (1998). The effect of cochlear implantation on vowel production. *Paediatric Cochlear Implantation. Abstract book 1998.*

Rietveld, T., Ebben, H., & **Huinck, W.** (1996). The influence of F_0 movements on the perception of the duration of silent intervals. *Proceedings 1996, department of Language and Speech, University of Nijmegen.*

Van de Velde, H., De Schutter, G., van Hout, R., Adank, P, **Huinck, W**., & Op 't Eynde, L., (1998). The pronunciation of standard Dutch in Flanders and the Netherlands. ESF sociolinguistic conference. *Convergence and divergence in a changing Europe. Poster presentation 16-19 september.*

Van Hout, R., De Schutter, G., De Crom, E., **Huinck, W.,** Kloots, H., & Van de Velde, H. (1998). De uitspraak van het Standaard-Nederlands. Variatie en varianten in Vlaanderen en Nederland. *Abstract book Lunteren.*

Van Hout, R., De Schutter, G., De Crom, E., **Huinck, W.,** Kloots, H., & Velde, H. Van de. (1999). De uitspraak van het Standaard-Nederlands: variatie en varianten in Vlaanderen en Nederland. In E. Huls & B. Weltens (Eds.), Artikelen van de Derde Sociolinguïstische Conferentie (pp. 183–196). Delft: Eburon.

Huinck, W.J., & Peters, H.F.M. (1999). Terminologie voor het beschrijven van stotteren. *Stem-, Spraak- en Taalpathologie, 8*(4), 210-219.

Huinck, W.J., Peters, H.F.M., & Hulstijn, W. (2000). The influence of gestural overlap during consonant transitions on stuttering. *Journal of Fluency Disorders, 25*(3), 174-175.

Huinck, W.J., Hulstijn, W., & Peters, H.F.M. (2000). The amount of Gestural Overlap in Consonant Clusters: Influences on the Fluent Speech of Persons Who do Stutter and Persons Who do Not Stutter. *ASHA Convention Program book 5*(16), p.167.

Huinck, W.J., Hulstijn, W., & Peters, H.F.M. (2001). The influence of Gestural Overlap during Consonant transition on Stuttering. *Proceedings IFA* page 72-77.

Huinck, W.J., Wouters, E.H.A., Hulstijn, W., & Peters, H.F.M. (2001). Differences between stuttering and non-stuttering people: comparing diadochokinesis, sentences repetition and reaction time tasks. In: B. Maassen, W. Hulstijn, R.D. Kent, H.F.M. Peters, & P.H.H.M. Van Lieshout (Eds.). *Speech motor control in normal and disordered speech* (pp. 311-315). Nijmegen, the Netherlands: Vantilt.

Diepstra, H.D., **Huinck, W.J.,** Hulstijn, W., & Peters, H.F.M. (2001). The effect of word length on speech reaction time and word duration in the fluent word production of stuttering and non-stuttering subjects. In: B. Maassen, W. Hulstijn, R.D. Kent, H.F.M. Peters, & P.H.H.M. Van Lieshout (Eds.). *Speech motor control in normal and disordered speech* (pp. 245-248). Nijmegen, the Netherlands: Vantilt.

Huinck, W.J. (2002). De evaluatie van drie verschillende stottertherapieën. Opzet en de eerste resultaten. *Proceedings Stem-, Spraak- en Taalstoornissen: onderzoek in ontwikkeling. Nijmegen*, 12 december 2002 (p.21-23).

Huinck, W.J., & Peters, H.F.M. (2004). Efficacy Research in Stuttering Therapy A longitudinal observation of the effects of three treatment programmes. Nijmegen: Nijmegen University Press.

Huinck, W.J., & Peters, H.F.M. (2004). Effectiviteitonderzoek bij stottertherapieën. Een longitudinale observationele studie naar therapeutische effecten van drie verschillende therapievormen. Nijmegen: Nijmegen University Press.

Huinck, W.J., van Lieshout, P.H.H.M., Peters, H.F.M., & Hulstijn, W. (2004). Gestural overlap in consonant clusters: effects on the fluent speech of stuttering and non-stuttering subjects. *Journal of Fluency Disorders, 29,* 3-25.

De Jong, F.I.C.R.S., Kooijman, P.G.C., **Huinck, W.J.**, Graamans, K., & Schutte, H.K., (2006). Epidemiology of voice problems in Dutch teachers. *Pholia Phoniatrica et Logopaedica, 58,* 186-198.

Kooijman, P.G.C., de Jong, F.I.C.R.S., Oudes, M.J., **Huinck, W.J.,** Acht, H. van, & Graamans, K. (2005). Muscular tension and body posture in teachers with persisting voice complaints. *Pholia Phoniatrica et Logopaedica, 57*, 134-147.

Kooijman, P.G.C., Jong, F.I.C.R.S. de, **Huinck, W.J.,** Graamans, K., & Schutte, H.K. (2006). Risk factors for voice problems in teachers. *Pholia Phoniatrica et Logopaedica, 58,* 159-174.

Huinck, W.J., & Donders, A.R.T. (2005). Meten van effectiviteit van stottertherapie. In H.F.M. Peters, R. Bastiaanse, J. Van Borsel, Ph. Dejonckere, K. Jansonius-Schultheiss, S. van der Meulen, and B.J.E. Mondelaers (Eds.), *Handboek Stem-, Spraak-, en Taalpathologie* [B7.3.3], (pp 1-18). Houten: Bohn Stafleu Van Loghum.

Rietveld, A., **Huinck, W.J.,** & Donders, A.R.T. (2005). Effectmeting in de taal- en spraakpathologie. In H.F.M. Peters, R. Bastiaanse, J. Van Borsel, Ph. Dejonckere, K. Jansonius-Schultheiss, S. van der Meulen, and B.J.E. Mondelaers (Eds.), *Handboek Stem-, Spraak-, en Taalpathologie* [A10.1], (pp 1-24). Houten: Bohn Stafleu Van Loghum.

Huinck, W.J., Langevin, M., Kully, D., Graamans, K., Peters, H.F.M., & Hulstijn, W. (2006). The relationship between pre-treatment clinical profile and treatment outcome in an integrated stuttering program. *Journal of Fluency Disorders 31*, 43-63.

Langevin, M., **Huinck, W.J.,** Kully, D., Peters, H.F.M., Lomheim, H., & Tellers, M. (2006). A Cross-Cultural, Long-Term Outcome Evaluation of the ISTAR Comprehensive Stuttering Program across Dutch and Canadian Adults who Stutter. *Journal of Fluency Disorders, in press.*

Huinck, W.J., & Rietveld, A.C.M. The validity of a simple outcome measure to assess stuttering therapy efficacy. *July 2006 accepted for publication in Pholia Phoniatrica et Logopaedica.*

Presentations

Huinck, W.J. (1999). Effectiviteitsonderzoek bij Stottertherapieën: Een observationele studie naar therapeutische effecten van drie verschillende therapievormen. *Nederlandse Federatie Stotteren, Utrecht.*

Huinck, W.J. (1999). Effectiviteitsonderzoek bij Stottertherapieën: Een observationele studie naar therapeutische effecten van drie verschillende therapievormen. Academisch Ziekenhuis Vrije Universiteit Amsterdam. *Nederlandse Vereniging Voor Stem-, Spraak- en Taalpathologie, Academisch Ziekenhuis Vrije Universiteit Amsterdam.*

Huinck, W.J. (2000). The influence of gestural overlap during consonant transitions on stuttering. IFA 3rd World Congress on Fluency Disorders. August 7-11, 2000 Nyborg, Denmark.

Huinck, W.J. (2000).The amount of Gestural Overlap in Consonant Clusters: Influences on the Fluent Speech of Persons Who do Stutter and Persons Who do Not Stutter. The American Speech-Language-Hearing Association (ASHA) annual convention. November 16-19, 2000 Washington DC, USA.

Huinck, W.J., Wouters, E.H.A., Hulstijn, W., & Peters, H.F.M. (2001). Differences between stuttering and non-stuttering people: comparing diadochokinesis, sentences repetition and reaction time tasks. 4th Speech Motor Conference, Nijmegen 2001.

Diepstra, H.D., **Huinck, W.J**., Hulstijn, W., & Peters, H.F.M. (2001). The effect of word length on speech reaction time and word duration in the fluent word production of stuttering and non-stuttering subjects. 4th Speech Motor Conference, Nijmegen 2001. Poster presentation.

Huinck, W.J., van Daalen, J., Hulstijn, W., & Peters, H.F.M. (2002). Variability on Respiratoy behavior in fluent and non-fluent speech. The American Speech-Language-Hearing Association (ASHA) annual convention. November, 2002 Atlanta, USA. Poster presentation.

Huinck, W.J., (2002). De evaluatie van drie verschillende stottertherapieën. Opzet en de eerste resultaten. Stem-, Spraak- en Taalstoornissen: onderzoek in ontwikkeling. Nijmegen12 december 2002.

de Jong, F.I.C.R.S., Kooijman, P.G.C., & **Huinck, W.J.** (2002). Voice problems in Dutch teachers: an epidemiological study Fourth congress of the European laryngological society, Brussels, 5–7 September 2002.

Huinck, W.J., & Peters, H.F.M. (2003). The effect of speech therapy on stuttering; Evaluating three therapy programs. *IFA conference Montreal August 2003.*

Huinck, W.J., & Peters, H.F.M. (2004). Efficacy in three therapy programmes: Van Riper Group therapy, The Comprehensive Stuttering Program and Individual Stuttering Therapy. 26th World Congress of the International Association of Logopedics and Phonetics. 29 august- 2 september 2004. Brisbane, Australia.

Kooijman, P.G.C., de Jong, F.I.C.R.S., Oudes, M.J., **Huinck, W.,** van Acht, W., & Graamans, K. Spierspanning en Lichaamshouding bij Leerkrachten met Stemklachten 204e Algemene KNO-vergadering, Nieuwegein.

Nijland, L., Maassen, B., & **Huinck, W.J.** (2004). Objectivering van spraakdiagnostiek. Dag van de Fonetiek. 23 december 2004.

Huinck, W.J., & Peters, H.F.M. (2006). *Efficacy Research in Stuttering Therapy: A Longitudinal Observation of the Effects of Three Treatment Programs* IFA conference Dublin July 2006.

Curriculum Vitae

Wendy Huinck behaalde in 1995 het diploma logopedie aan de Hogeschool Heerlen, waarna ze de opleiding Spraak- en Taalpathologie aan de Katholieke Universiteit van Nijmegen volgde. Naast deze opleiding werkte ze drie dagen in de week. Eerst als logopedist, daarna als onderzoeksassistent bij het cochleaire implantatieteam in Utrecht. In januari 1998 studeerde ze af als spraak- taalpatholoog. Aansluitend werkte ze vier dagen in de week als onderzoeker op het onderzoeksproject "De uitspraak van het Standaardnederlands: variatie en varianten in Nederland en Vlaanderen". Eén dag in de week bleef ze werkzaam als onderzoeksassistent binnen het cochleaire implantatieteam in Utrecht. In januari 1999 maakte ze de overstap naar het onderzoeksproject "Effectiviteitonderzoek bij stottertherapieën" gefinancierd door het College van Zorgverzekeringen, waarbij het effect van drie verschillende stottertherapieën werd onderzocht. Eind 2004 eindigde dit onderzoeksproject en dit heeft geresulteerd in dit proefschrift. Sinds november 2005 is ze aangesteld als coördinator cochleaire implantie bij volwassenen binnen het cochleaire implantatieteam verbonden aan de afdeling KNO van het Universitair Medisch Centrum St Radboud te Nijmegen.